A TIME-TO-FIRST-SPIKE CMOS IMAGE SENSOR

By

XIN QI

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2004

To my parents and my wife

ACKNOWLEDGMENTS

I would like to express my sincere gratitude to my advisor, Dr. John G. Harris, for his support, critique and encouragement in the past four years. Without his patience and guidance, none of this work would be possible. I would like to thank Dr. Robert M. Fox for his teaching on analog circuit design and Dr. William R. Eisenstadt on VLSI design. I also thank Dr. Stephen S. Eikenberry for serving on my graduate committee. Throughout my PhD research, Dr. Xiaochuan Guo stands out in helping me understand CMOS image sensor design. I am very grateful for his kindness and guidance. I also owe many thanks to my colleagues in the Computational Neuroengineering Lab and my friends in the Electrical and Computer Engineering Department for their discussions of ideas and friendships.

I would like to express my deepest appreciation to my wife, Dong Guo, for her great love and support. She brings me so much happiness and converts my "tedious" study period into an enjoyable life. Finally, I am especially grateful to my parents in China for their love and support. I dedicate this dissertation to them.

TABLE OF CONTENTS

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

A TIME-TO-FIRST-SPIKE CMOS IMAGE SENSOR

By

Xin Qi

December 2004

Chair: John G. Harris
Major Department: Electrical and Computer Engineering

Complementary metal oxide semiconductor (CMOS) image sensors have become more attractive recently due to their capabilities of offering system-on-chip (SoC), low cost and low power consumption. Conventional CMOS image sensors, however, have a limited dynamic range of 65-75dB, while a typical outdoor scene has a dynamic range of more than 100dB. The purpose of this dissertation is to design a novel CMOS image sensor to deal with such high dynamic range scenes. The dissertation begins with a review of basic physics of photodetectors. The following analysis of conventional CMOS imager principles reveals that a limited output signal swing and a fixed integration time result in the dynamic range limitation. To circumvent this problem, we proposed a novel time-to-first-spike (TTFS) CMOS imager, which encodes illuminance with integration time instead of reading out the analog value at a fixed integration time. The principle and circuit design of this novel TTFS imager are discussed in detail. Our analysis has shown that this TTFS imager is able to achieve high dynamic range as well as good signal-to-noise ratio performance. A prototype chip was fabricated in TSMC 0.18μm digital technology and fully tested. The testing results verify the expected performance. Some optimization techniques, e.g., optimal

variation of reference voltage, and several modified TTFS imager designs are also presented in this dissertation.

CHAPTER 1
INTRODUCTION

Invented in 1970, charge-coupled devices (CCDs) are presently the technology of choice for most imaging applications due to their high sensitivity, high quantum efficiency and large fill factor [1]. The demand for near-perfect charge transfer efficiency, however, makes CCDs difficult to scale up to very large array size and achieve high readout speed. Fabricated in a special process, CCDs are incapable of integrating other electronic circuits on the same chip. In addition, they need high voltage operations, thus consuming more power [2].

The fast growing digital multimedia application market requires miniaturized, low cost and low power consumption cameras. This demand has attracted more attention to the design of CMOS-based image sensors. CMOS image sensors are fabricated in standard CMOS technologies and potentially able to integrate a significant amount of VLSI electronics, e.g., timing control, analog-to-digital converters (ADC) and signal processing, into a single chip. Therefore, they greatly reduce component and package cost. Powered by standard logic supply voltages, CMOS image sensors also benefit from very low power consumption measured in the tens of milliwatts for a 256×256 size array [3].

One serious problem with CMOS image sensors is the limited dynamic range (DR). Dynamic range is typically defined as the ratio of the largest nonsaturated signal to the smallest measurable signal. Conventional CMOS image sensors are usually limited to 65-75dB DR due to the narrow signal swing and a single integration time for all pixels, while a typical outdoor scene has a dynamic range of more than 100dB. With standard CMOS technology scaling down to submicron levels, DR tends to be further deteriorated because of the reduced signal headroom and increased noise

floor. Inspired by biological vision theory, we have proposed a time-to-first-spike (TTFS) imager [4]. Instead of reading out analog signals, we encode the illuminance intensities with temporal information. There is not a fixed integration time as with conventional CMOS imagers; however the integration time of each pixel varies with respect to the illuminance. Thus it overcomes the restriction of limited voltage swing and extends the dynamic range in the time domain. The early version of the TTFS imager had a large pixel size and a potentially high collision rate for a large size array. In this dissertation, we will optimize our original architecture at both system and circuit levels. The dissertation is organized as follows: Chapter 2 provides the fundamentals of solid-state image sensors. In Chapter 3, we will review the existing high dynamic range image sensors and describe the principles of our novel CMOS image sensor. An optimal reference voltage variation strategy for TTFS imagers is presented in Chapter 4. Chapter 5 describes the circuit design and presents testing results. Several modified TTFS imagers with low readout collision rate are included in Chapter 6. Finally, the dissertation is concluded with the future work in Chapter 7.

CHAPTER 2
FUNDAMENTALS OF SOLID-STATE IMAGE SENSORS

A solid-state image sensor, an integrated circuit that contains a number of photodetectors in a 2D array, converts an optical signal into an electrical output. A basic physics review of photodetectors is provided in this chapter. Following that, we will briefly introduce the two most important solid-state imagers: CCD and CMOS image sensors. A noise analysis is given at the end of this chapter.

2.1 Physics Review of Solid-State Photodetectors

Photodetectors are solid-state devices that detect optical signals and convert them into electrical signals [5]. This photodetection process can be summarized with the following steps:

1. Absorption of photons to generate charge carriers,
2. Drift of charge carriers in a built-in electric field,
3. Collection of charge carriers.

We will explain these steps in the following sections.

2.1.1 Generation of Charge Carriers

If the incident photons have energies greater than the bandgap of semiconductor photodetectors, electron-hole (e-h) pairs are generated inside the semiconductors. Then the electrons in the valence band are brought to the conduction band while leaving holes behind. According to Planck's relationship, the photon energy is

$$E_{ph} = h\nu = h\frac{c}{\lambda} \tag{2.1}$$

where $h = 6.626\times10^{-34} J\text{-}s$ is the Planck constant, $c = 3\times10^{8} m/s$ is the speed of light, and λ is the wavelength. To generate the e-h pairs, the wavelength of incident light should be shorter than $\lambda_s[\mu m] = hc/E_g = 1.24/E_g[eV]$. For silicon, $E_g = 1.12eV$,

the threshold wavelength is $\lambda_s \approx 1110nm$, whereas for Ge with $E_g = 0.66eV$, the corresponding wavelength is $\lambda_s \approx 1880nm$. Since both these threshold wavelengths are longer than the visible light wavelength that ranges from about $400nm$ (violet) to $700nm$ (red) [6], both silicon and germanium photodetectors can be used to sense visible light.

An absorbing material, e.g., silicon, generally absorbs some of the energy of the incident light and reflects the rest. For simplicity, we assume here that all the incident light is available for e-h pair generation and no reflection occurs, i.e., $R = \frac{reflected\ intensity}{incident\ intensity} = 0$. Then, how much energy absorbed is decided by the optical absorption coefficient α, a function of wavelength. Related to the absorption coefficient, the light intensity $F(x)$ at the depth of x is given by Beer's law

$$F(x) = F_0 e^{-\alpha x} \tag{2.2}$$

where F_0 is the light intensity at the surface. Beer's law states that the light intensity exponentially decreases inside the absorbing material. So alternatively, the absorption coefficient is determined by the penetration depth of the light, which is defined as the location where the light intensity becomes $1/e$ (63%) of the surface light intensity, i.e., $D_{pen} = 1/\alpha$. Typical absorption coefficients of silicon shown in figure 2–1 indicate that the penetration depths of violet light and red light are approximately $D_{pen} = 0.17\mu m$ and $D_{pen} = 40\mu m$ respectively. This information can be used as the basis for color sensors, e.g., the Foveon X3 digital image sensor [7].

2.1.2 Photon Collection and Quantum Efficiency

The reverse process of *e-h* generation in a semiconductor is recombination, which can annihilate *e-h* pairs generated by the incident light. Fortunately, we can use an electrical field to effectively separate electrons and holes and cause the carriers to reach some collection contacts [8]. Thus photocurrent, J_{ph}, is formed, which has three components:

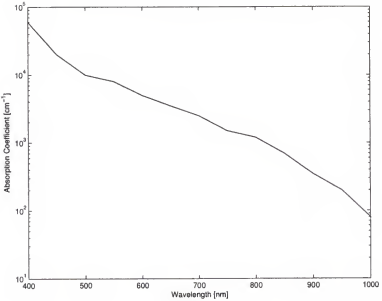


Figure 2–1: Absorption coefficients of silicon (after Gamal [6])

1. Current due to the carriers generated in the depletion region and swept away by a strong electric field, J_{drift},

2. Current due to holes (minority carrier) generated in the n-type quasi-neutral region, J^n_{diff},

3. Current due to electrons (minority carrier) generated in the p-type quasi-neutral region, J^p_{diff}.

Figure 2–2 illustrates the photocurrent generation mechanism in a pn junction photodiode. Assuming the n-layer is thin enough to cause negligible absorption, i.e., $x_1 = 0$ [9], and the p-type region is the lower concentration part of the junction, then the total current density through the depletion layer is

$$J_{ph} = J_{drift} + J^p_{diff} \tag{2.3}$$

Assuming a monochromatic incident light with intensity F_0 at the surface, the carrier generation rate at the depth of x is given by

$$G(x) = \frac{d}{dx}\left(F_0 - F(x)\right) = \alpha F_0 e^{-\alpha x} \tag{2.4}$$

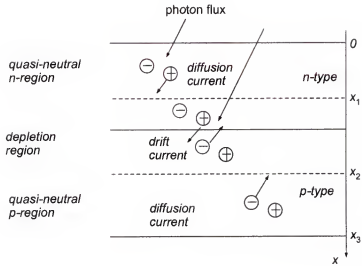


Figure 2–2: Diagram of photocurrent generation

If an abrupt pn junction is assumed and all the generated carriers in the depletion layer are collected, the drift current density is given by

$$J_{drift} = \int_0^{x_2} qG(x)dx = qF_0(1 - e^{-\alpha W}) \tag{2.5}$$

where $q = 1.6 \times 10^{-19}C$ is the electron charge, and $W = x_2$ is the depletion width as labelled in figure 2–2.

Unlike the drift current, the diffusion current in the quasi-p region is governed by the one-dimensional diffusion equation

$$D_n \frac{\partial^2 n_p}{\partial x^2} - \frac{n_p - n_{p0}}{\tau_n} + G(x) = 0 \tag{2.6}$$

where D_n is the diffusion coefficient for electrons, τ_n is the lifetime of excess carriers, and n_{p0} is the equilibrium electron density. With the boundary conditions $n_p = 0$ @$x = W$ and $n_p = n_{p0}$ @$x = \infty$, we can solve the equation to get

$$n_p = n_{p0} - \left(n_{p0} + C_1 e^{-\alpha W}\right) e^{(W-x)/L_n} + C_1 e^{-\alpha x} \tag{2.7}$$

where $L_n = \sqrt{D_n \tau_n}$ and $C_1 \equiv \left(\frac{F_0}{D_n}\right) \frac{\alpha L_n^2}{1-\alpha^2 L_n^2}$. Now the diffusion current density is given by

$$\begin{aligned} J_{diff}^{p} &= qD_n \left(\frac{\partial n_p}{\partial x}\right)_{x=W} \\ &= qF_0 \frac{\alpha L_n}{1+\alpha L_n} e^{-\alpha W} + qn_{p0}\frac{D_n}{L_n} \end{aligned} \tag{2.8}$$

Combining equation 2.5 and 2.8, we obtain the total current density

$$J_{ph} = qF_0 \left(1 - \frac{e^{-\alpha W}}{1+\alpha L_n}\right) + qn_{p0}\frac{D_n}{L_n} \tag{2.9}$$

As a rule of thumb, it is best to widen the depletion region to absorb more photons [10]. For a one-sided silicon pn diode, the depletion region width can be described by

$$W = \sqrt{\frac{2\varepsilon_{si}(\psi_{bi} + V_b)}{qN_b}} \tag{2.10}$$

where ε_{si} is the permittivity of silicon, ψ_{bi} is the built-in potential of a one-sided pn diode, V_b is the externally applied reverse bias to the junction, and N_b is the lightly doped density, either the donor or the acceptor concentration. According to the above equation, the depletion region can be widened by lowering the doping density or increasing the reverse bias to the junction.

In applications, the photodetector response is characterized by quantum efficiency (QE) $\eta(\lambda)$, defined as the ratio of the collected charge pairs to the absorbed photons. Ignoring the second term of equation 2.9, we have

$$\eta(\lambda) = 1 - \frac{e^{-\alpha W}}{1+\alpha L_n} \tag{2.11}$$

Similar to the incident light generated photocurrent, QE is also a function of wavelength and directly related to the process parameters via L_n and W.

In a typical CMOS process, three types of vertical P/N diodes, i.e., n+/psub, p+/nwell and nwell/psub, are available. The reported QEs for n+/psub and p+/nwell

diodes are lower than that for nwell/psub, which is mainly due to two reasons. First, silicide technology, widely used over the n+/p+ regions in today's advanced technologies [11], is opaque to visible light and degrades the light sensitivity. In most technologies, it is not feasible to access the device process with a silicide block. Second, a nwell/psub diode has a wider depletion region compared to the other two diodes under the same bias condition, thus collecting more carriers.

2.1.3 Dark Current

Dark current is the thermal generated photodiode leakage current, which does not depend on the incident light intensity. It comes from the defects either at the SiO_2-Si surface or in the bulk. Among them, the high irregularity at the SiO_2-Si surface is the principal contributor of dark current, which has been verified by the measured data in Loukanova et al. [12]. Dark current is an important parameter to characterize the performance of image sensors since the variations of dark current from pixel to pixel will contribute to the fixed pattern noise (FPN). In addition, the shot noise associated with dark current is part of the total temporal noise, which could deteriorate the signal-to-noise ratio (SNR) and dynamic range (DR) performance of image sensors. So it is preferred to design image sensors with a low dark current.

Under low-biased conditions, the dark (leakage) current of a pn junction consists of the diffusion current from the quasi-neutral areas and the thermal generation current from the depletion region. Adopting the same device structure described in the previous section, the dark current density can be expressed as [9]

$$\begin{aligned} J_{dark} &\approx J_{diff}^{dark} + J_{SRH}^{dark} \\ &= q\sqrt{\frac{D_n}{\tau_n}}\frac{n_i^2}{N_A} + q\frac{n_i}{\tau_{gen}}W \end{aligned} \tag{2.12}$$

where D_n is the electron diffusivity, τ_n is the electron life time, N_A is the doping level of p-type region, n_i is the intrinsic concentration of silicon, τ_{gen} is the generation lifetime, and W is the depletion width. The foregoing derivation reveals that the

dark current is a strong function of the depletion region width and the temperature. The temperature dependence mainly comes from the intrinsic concentration of silicon, n_i [12]. Generally, a high temperature results in a large intrinsic concentration, thus a high dark current.

Physically, there are two feasible ways to reduce the dark current, i.e., narrowing the depletion region or decreasing the working temperature. Since a narrow depletion region may degrade the quantum efficiency, a low working temperature becomes the best choice. In addition, many other efforts have been tried to lower dark current through either an optimized layout approach or a clever readout circuit design. Since the SiO_2-Si surface defect is the dominant dark current source, a small dark current can be achieved if the photon-sensing area is isolated from the defective field oxide edge. One way is to modify the process with a shallow active layer above the photon-sensing area to form a new structure, called pinned photodiode [13]. Another way is to employ a reset gate-poly ring surrounding the photodiode, which can eliminate the dark current originating from the border region adjacent to the defective field oxide edge without modifying the process [14]. Aside from optimal pixel designs, a novel readout scheme with a combined photogate/photodiode photo-sensing device is also able to achieve an ultra-low dark current. Thanks to the correlation of dark currents in neighboring readout signals, the dark currents can be somewhat cancelled in the output signal by differentiating two neighboring readout signals [15].

2.2 Charge-Coupled Devices (CCDs)

The CCD was invented in 1970. This dynamic (charge) shift register is implemented using closely spaced MOS capacitors with 2, 3 or 4 phase clocks (see figure 2–3). CCDs are optimized photodetectors, whose virtues include high QE, low dark current, low noise and high fill factor.

A CCD imager generally adopts a serial readout technique as illustrated in figure 2–4. The analog charge must be shifted out of the chip before it converts into a

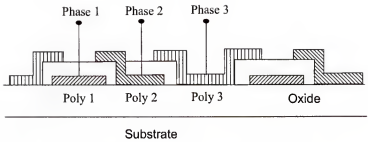


Figure 2–3: A three phase CCD (after Gamal [6])

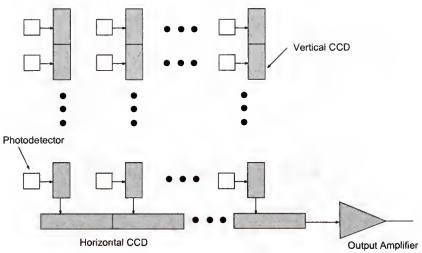


Figure 2–4: An interline transfer CCD image sensor (after Gamal [6])

digital value. It requires, on average, several thousands of shifts for a large CCD array. Therefore the charge transfer efficiency (CTE), defined as the fraction of signal charge transferred from one CCD stage to the next, should be high enough to avoid charge loss. To understand this serious issue, let us consider a CTE given by $\eta = 0.999$. Then the net fraction of signal transferred after $m = 1024$ stages is only $\eta^m = 0.359$, which means a great amount of charge has been lost during the transfer process. The need for near-perfect CTE has a great impact on CCD imagers. In summary, the major reported limitations of CCDs are [6, 16]:

1. Limited frame rate due to the minimum required transfer time per CCD stage to ensure a perfect CTE,

2. Difficulty to achieve a large size array limited by the CTE,

3. Requiring complex high speed shifting clocks for proper operations,

4. High power consumption since a relative high voltage (up to 15V) is required for proper operations and the entire array are switching all the time,

5. Highly nonprogrammable,

6. Inability to implement system-on-chip (SoC).

Though CCDs still dominate in the commercial products market, the limitations have driven researchers to study new solutions to circumvent the major weaknesses of CCD technology. CMOS image sensors emerge at this request. We will address this technology in the next section.

2.3 CMOS Image Sensors

A CMOS image sensor, fabricated in a standard CMOS process, is capable of integrating timing and control electronics, a sensor array, signal processing electronics, an analog-to-digital converter (ADC) and a full digital interface on the same chip [1]. It operates with standard logic supply voltages and consumes little power. Recent advances have made CMOS image sensors' performance competitive with CCDs. Most CMOS image sensors are two-dimensional addressable arrays as shown in figure 2–5. After integrating photocurrent for a predefined period, the analog charge is read out by transferring one row at a time to the column sample-and-hold circuits, then reading out one or more pixels in the selected row at a time using the multiplexer.

2.3.1 A Simple Equivalent Circuit of Photodiode

As discussed, most CMOS image sensors operate in the integration mode, and the photocurrent is read out in terms of integrated charge. Since photodiodes are widely used in CMOS image sensors, a simple equivalent circuit of a photodiode is needed for the purpose of analysis and simulation. The equivalent circuit model of a photodiode is shown in figure 2–6. It is made up of a current source and a capacitor in parallel. Once it is reset, the capacitor is discharged from the initial reset voltage V_{reset} by the

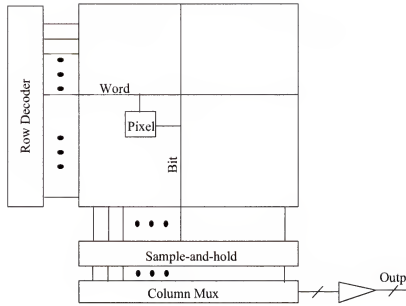


Figure 2–5: A typical CMOS image sensor architecture

normalized photocurrent $I(V_{out})$ given in equation 2.9. Hence, the following relation holds

$$\int_{V}^{V_{reset}} \frac{C_{pd}(V_{out})}{I(V_{out})}\, dV_{out}(t) = \int_{0}^{T} dt \tag{2.13}$$

where $C_{pd}(V_{out}) = C_0 + C_d(V_{out})$. C_0 is the peripheral capacitance normalized by the photodiode junction area A, and $C_d(V_{out})$ is the unit depletion capacitance of the photodiode and a function of V_{out}, $C_d(V_{out}) = \frac{\varepsilon_{si}}{W_d(V_{out})}$. It is very difficult to derive a closed form solution for the above equation. Here we only provide the numerical simulation results for a nwell/psub diode with $V_{reset} = 3.3$V in figure 2–7 and figure 2–8.

Two scenarios are considered here. First, the peripheral capacitance C_0 is negligible. Second, C_0 is much larger than C_d. Surprisingly, both cases demonstrate good linearity for integrated charge with respect to incident light intensity and dropping output voltage with respect to integration time. So the photodiode model can be simplified to a constant capacitor C_{pd} and a constant photocurrent source I in parallel.

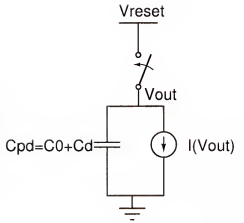


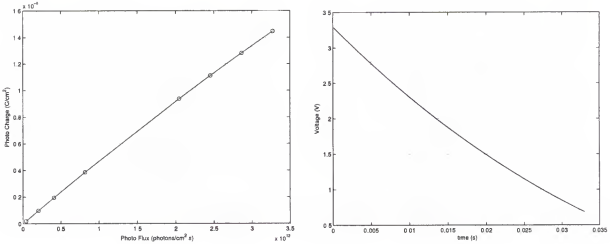
Figure 2–6: An equivalent model of photodiode

2.3.2 Pixel Circuits

Conventional CMOS image sensors can be divided into two categories: passive pixel sensors (PPS) and active pixel sensors (APS). The PPS concept is shown in figure 2–9, which consists of a photodiode and an access switch. When the photodiode is accessed, the integrated charge is converted to a voltage by a charge integrating amplifier located at the bottom of the column bus. The large capacitive load causes very slow readout speed. Fortunately, with the insertion of a buffer into the pixel, an APS sensor can efficiently solve the readout speed problem with PPS. Figure 2–10 shows a typical active pixel schematic. In steady state by assuming charge Q accumulated on the photodiode capacitance C_{pd} at the end of integration and ignoring the voltage drop across the access transistor and body effect, we obtain the output voltage

$$V_o = V_{dd} - \frac{Q}{C_{pd}} - V_{th} - \sqrt{\frac{2L_F}{C_{ox}\mu_n W_F} I_{bias}} \tag{2.14}$$

where V_{th} is the source follower threshold voltage, C_{ox} is the unit oxide capacitance, μ_n is the electron mobility, and W_F/L_F is the source follower transistor size. Obviously, the readout voltage V_o directly reflects the integrated photosignal.



(a) (b)

Figure 2–7: Simulated photodiode response when C_0 is negligible. (a) integrated charge vs. incident light intensity for wavelenth=500nm. (b) output voltage vs. integration time for $F_0 = 3.2 \times 10^{12} photons/cm^2 s$ at the surface.

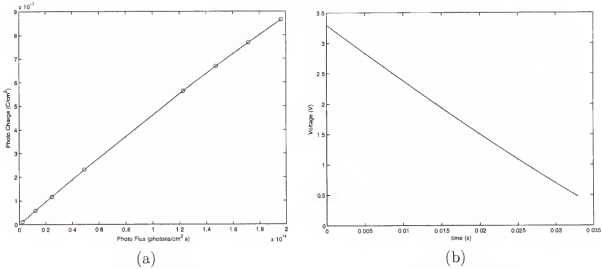
2.4 Noise in Image Sensors

Like other electronic products, the overall performance of an image sensor is ultimately determined by the noise introduced by the system into the signal. Noise in image sensors is typically divided into temporal noise and fixed pattern noise. We will address these noise sources separately in the following.

2.4.1 Temporal Noise

Temporal noise is the time-dependant fluctuations in the signal level due to device noise. It can be introduced from the pixel, the readout circuit, the substrate and the power supply. However, we will not include the circuit-oriented temporal noise originating from the substrate coupling or the power supply oscillation in the following discussion, because they are considered to be negligible compared to other noise sources.

1. Pixel Photon Shot Noise



(a) (b)

Figure 2–8: Simulated photodiode response when C_0 is much larger than C_d. (a) integrated charge vs. incident light intensity for wavelenth=500nm. (b) output voltage vs. integration time for $F_0 = 1.96 \times 10^{14} photons/cm^2 s$ at the surface.

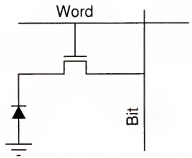


Figure 2–9: Passive pixel schematic

Photon shot noise[1] is essentially the result of the random generation of carriers and obeys the Poisson statistics. It is generated either by the thermal generation within a depletion region or by the random arrival of photons. The noise is expressed as

$$n_{photon} = \sqrt{\frac{Q_{photon}}{q}} = \sqrt{\frac{i_{ph} t_{int}}{q}} \quad (electrons)^2 \tag{2.15}$$

[1] In this dissertation, we also call it photocurrent shot noise.

[2] In this dissertation, we also use e^- to represent *electron*.

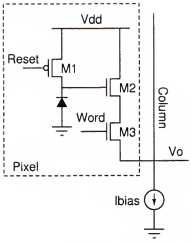


Figure 2–10: Active pixel schematic

where Q_{photon} is the integrated charge introduced by photons, i_{ph} is the generated photocurrent, and t_{int} is the integration time.

2. Pixel Dark Current Shot Noise

   Same as the photon shot noise, the thermal generation of charge carriers under dark conditions is also a Poisson process. If Q_{dark} is the integrated charge due to a dark current i_{dark} within an integration period of t_{int}, the dark current shot noise is given by

$$n_{dark} = \sqrt{\frac{Q_{dark}}{q}} = \sqrt{\frac{i_{dark} t_{int}}{q}} \quad (electrons) \tag{2.16}$$

3. Pixel Reset (kT/C) Noise

   Introduced by the channel thermal noise of the reset transistor (i.e., $M1$ in figure 2–10), the reset noise on the photodiode is given by[3] [17]

$$\overline{V_{rst}^2} = \frac{kT}{C_{pd}} \tag{2.17}$$

[3] By using a NMOS as the reset transistor, the reset noise power is $\overline{V_{rst}^2} = \frac{kT}{2C_{pd}}$ if the steady state is not reached during reset.

where $k = 1.38 \times 10^{-23} J/K$ is the Boltzmann constant, and T is the temperature in Kelvin.

4. In-Pixel MOS Device Noise

   The in-pixel amplifier noise primarily comes from the thermal and flicker noise of MOS transistors. They may be suppressed by a DC offset cancellation technique or limiting the bandwidth of the amplifier.

5. Readout Noise

   The MOS transistors in the column-level or chip-level readout circuit will inevitably introduce thermal noise and flicker noise, which are independent of photocurrent, dark current and integration time.

2.4.2 Fixed Pattern Noise (FPN)

Fixed pattern noise is a non-temporal spatial noise, which is due to device mismatches in pixels and column-level circuits. The major components include:

1. Dark current FPN due to the mismatch in photodiode leakage currents.
2. Pixel response FPN due to the nonuniformities of geometry, layer thickness or doping profile.
3. Readout circuit FPN mainly due to the threshold voltage variations between MOSFETS.

With the application of a double-delta-sampling (DDS) circuit [3], the readout circuit FPN can be suppressed to a negligible level compared to the photocurrent FPNs, i.e., the dark current FPN and the pixel response FPN. Under this condition, the FPN is proved to be dominated by the dark current FPN at low signal levels and the pixel response FPN at high signal levels [18].

2.5 Discussion

The previous derivations of photocurrent and quantum efficiency are very simple but instructive. A more explicit analysis may include the diffusion current in the quasi-n region, the reflection at the surface of the chip, the reflections and absorptions

in the layers above the photodetectors and the variation of J_{ph} over the photodetector area (edge effects) [6]. Though the physics of photodetectors reviewed in this chapter is quite fundamental, a good understanding is beneficial to our system design.

CHAPTER 3
REVIEW OF HIGH DYNAMIC RANGE CMOS IMAGERS

A typical CMOS APS has a dynamic range of 65-75dB [19]. To improve the dynamic range, two approaches are considered. One is to reduce the noise floor and extend the dynamic range towards the weak signal region. The other is to expand the nonsaturating signal level. In this dissertation, we only concentrate on the latter. This chapter starts with the analysis of the dynamic range problem with conventional CMOS APSs, followed by a review of existing dynamic range enhancement methods. Finally, our novel biologically inspired time-to-first-spike imager is presented.

3.1 DR Problem with Conventional CMOS APSs

In this section, we take the most commonly used photodiode APS as an example to show the DR limitation. For an APS imager shown in figure 3–1, if $M4$ operates in the strong inversion region and saturates, the output voltage in terms of the integrated charge is

$$V_o = V_{dd} - \frac{Q}{C_{pd}} - V_{th,F} - \sqrt{\frac{L_F}{W_F}\frac{W_4}{L_4}}(V_{bias} - V_{th,B}) \tag{3.1}$$

where $V_{th,F}$ is the source follower threshold voltage without considering body effect, $V_{th,B}$ is the bias transistor threshold voltage, and W_F/L_F, W_4/L_4 are the sizes of $M2$ and $M4$ respectively. The maximum output voltage occurs when $Q = 0$, which is

$$V_{omax} = V_{dd} - V_{th,F} - \sqrt{\frac{L_F}{W_F}\frac{W_4}{L_4}}(V_{bias} - V_{th,B}) \tag{3.2}$$

whereas the minimum output voltage V_{omin} is decided by the operation of the bias transistor $M4$. To ensure that $M4$ works in saturation, V_{omin} must be

$$V_{omin} = V_{bias} - V_{th,B} \tag{3.3}$$

It gives a signal swing of

$$\begin{aligned} V_s &= V_{omax} - V_{omin} \\ &= V_{dd} - V_{th,F} - \left(\sqrt{\frac{L_F}{W_F}\frac{W_4}{L_4}} + 1 \right) (V_{bias} - V_{th,B}) \end{aligned} \tag{3.4}$$

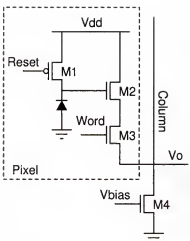


Figure 3–1: DR limitation of CMOS APS

Usually, we use well capacity instead of signal swing to evaluate the performance of CMOS image sensors. For a CMOS APS, the effective well capacity, $Q_{eff} = V_s \times C_{pd}$, is much less than the available maximum well capacity $Q_{max} = V_{dd} \times C_{pd}$, since the output voltage reaches its minimum value before the diode voltage drops to ground. If the dark current is negligible compared to the regular photocurrent, the largest nonsaturating signal can be computed as $i_{max} = Q_{eff}/t_{int}$ for an integration period of t_{int}. Generally, the smallest detectable input signal is defined as the input referred noise floor under dark conditions, which gives

$$i_{min} = \frac{q}{t_{int}} \sqrt{\frac{i_{dark} t_{int}}{q} + \sigma_r^2} \tag{3.5}$$

Here, i_{dark} is the dark current, and σ_r^2 $(electrons^2)$ includes the readout noise and the reset noise. As the ratio of i_{max} to i_{min}, the dynamic range equals

$$DR = 20 \log \frac{Q_{eff}}{\sqrt{i_{dark} t_{int} q + \sigma_r^2 q^2}} \quad (dB) \tag{3.6}$$

Equation 3.6 is plotted in figure 3–2 for a sensor with $V_s = 1\text{V}$, $C_{pd} = 10fF$, $\sigma_r = 20e^-$, and $i_{dark} = 1fA$. As the integration time increases, the increased dark current shot noise increases the noise floor monotonically and reduces the dynamic range. Note that whatever the integration time is, the maximum achievable DR is less than 70dB as a result of limited well capacity.

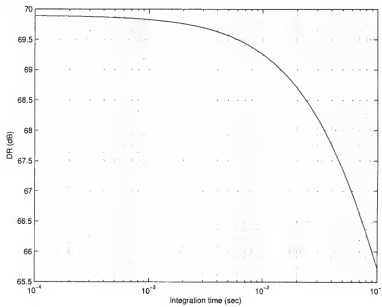


Figure 3–2: DR vs. integration time

The achievable dynamic range will be further deteriorated with CMOS technology continuously scaling down because of the reduced signal headroom and the increased noise floor. Figure 3–3 illustrates this trend with $t_{int} = 33ms$, $C_{pd} = 10fF$, $\sigma_r = 20e^-$, and $i_{dark} = 1fA$.

A typical outdoor scene has a DR of more than 100dB, which is far beyond conventional CMOS APSs' DR. In order to deal with such high dynamic range scenes, special designs are demanded. Current dynamic range enhancement methods will be addressed in the next section.

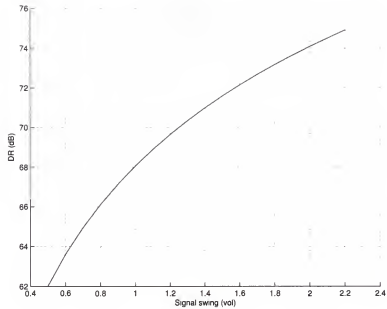


Figure 3–3: DR vs. signal swing

3.2 Existing Dynamic Range Enhancement Methods

3.2.1 Analog Domain Methods

All the methods in this category represent illuminance in the form of an analog signal, e.g., current, voltage or integrated charge. Another representation of illuminance will be discussed later.

Nonlinear Signal Compression (NSC)

As explained earlier, the limited output voltage swing and the linear mapping between output voltage and the incident light generated photocurrent limit the achievable DR to only 70dB. Since the voltage supply cannot be significantly increased, many researchers have tried to use nonlinear response curves to improve the dynamic range.

If the photocurrent is fed into a diode-connected MOS transistor in the subthreshold region, the output voltage will exhibit a logarithmic response of the input current. This nonlinear response voltage can cover about five orders of the magnitude

of the incoming light intensity or 100dB dynamic range. This technique has been successfully implemented by a research group at Caltech [20]. However, low contrast, loss of details, low SNR and high nonlinear FPN make it very problematic [19].

Compared with the logarithmic response imager, a well capacity adjustment technique has better control over the FPN. It uses a lateral overflow gate to compress the sensor's current versus charge response curve. A correlated double sampling (CDS) technique is adopted to reduce the FPN. The reported dynamic range is 96dB and the FPN is only 0.24% of saturation level [21]. However, this technique exhibits worse SNR performance [22].

Multiple Sampling (MS)

Conventional CMOS APSs have only one predefined integration time t_{int}. A long t_{int} favors weak signals by integrating much more charge but easily saturates strong signals, whereas a short t_{int} can prevent signals from saturating but degrades the SNR of weak signals. No single integration time can satisfy both dynamic range and SNR requirements. To solve this problem, a multiple sampling technique has been proposed, which uses shorter exposure times to capture the brighter parts of the scene and longer exposure times to capture the darker regions [22, 23]. The reported CMOS image sensors have shown the expected dynamic range and offered good image quality as well. This type of image sensor is very promising but fundamentally consumes more power and requires a large data bandwidth to readout the multiple frames.

3.2.2 Timed-Based Methods

Instead of choosing a single integration time as in conventional CMOS APSs, time-based image sensors allow each pixel to choose its own optimal integration time with respect to the illuminance. In this way, illuminance is encoded with temporal information much like neural coding in a biological vision system. Basically, time-based image sensors operate in two modes: first, asynchronous mode and second, synchronous mode.

Asynchronous Mode (TBAM)

The time-based CMOS image sensors working in the asynchronous mode have no global reset signal for all pixels. Instead, each pixel works as a free-running continuous oscillator and the whole system is essentially a frequency modulator in that high frequency pulses represent bright illuminance and low frequency pulses represent dark illuminance. Many different readout architectures have been implemented to sample the pulses. In the first time-based image sensor, an asynchronous counter on each row counts the output pulses for a fixed time period [24]. The short-term average pulse frequency is used to reconstruct the illuminance. A dynamic range of 120dB was achieved for a static scene.

By noticing the equivalence between a synchronous first-order Σ-Δ modulator and a sampled asynchronously running oscillator, McIlrath [25] uses the binary output stream from the sampled oscillator to reconstruct the scene. The author claims that by sweeping through a set of binary weighted frequencies, $f_s = f_0, f_0/2, \ldots, f_0/2^k$, only $8k$ samples need to be taken to give a dynamic range of $6k + 25$ dB.

The address-event circuit, originally developed to communicate spike trains between arrays of silicon neurons on multiple chips [26], can be applied in an imager. It works as follows: when a pixel reaches a threshold, a request (spike) for access to the output bus is sent out to the address-event circuit. Once the request is approved, the address of the pixel is readout. The average interspike interval is measured to represent the light intensity information [27].

Although the reconstruction methods vary, all the asynchronous mode time-based imagers in the literature have to readout a large amount of redundant information, which implies more power consumption, larger data bandwidth and more frame memory. One serious problem with these designs is that a long frame time may be needed to collect all useful information for the scene recovery, which is obviously not feasible for dynamic scenes or video mode applications.

Synchronous Mode (TBSM)

For synchronous mode time-based CMOS image sensors, there is a global reset signal in each imager. After a photodetector is reset, the voltage on the photodetector linearly decreases. Once the voltage drops below the threshold voltage, a time stamp is stored on a capacitor inside each pixel [28, 29, 30]. Since the eventual readout signal is an analog signal, it still needs an external ADC to obtain digital values.

In [31], a time domain sampling method is proposed. 2^k samples are required to get k bits of resolution, which means requiring much more memory and more power consumption.

3.3 Time-to-First-Spike Imager

Recently, some biologists have claimed that the most useful information from the retina is contained in the first spike after onset of the neuronal response. Based on psychophysical experiments, they have determined that reactions times are so short that there is no enough time to process more than one spike from each neuron per processing stage [32, 33]. Inspired by both the biology and engineering constraints, we propose a novel imager, a time-to-first-spike (TTFS) imager, which represents the illuminance in a different way. To overcome the shortcomings of the previous time-based designs, the illuminance is transformed into a pulse event that can *only occur once per pixel per frame* in the time domain. With respect to the imager's global (or row-level) reset signal that occurs at the start of each frame, a brighter pixel's pulse event occurs before a darker pixel's pulse event. The spike readout circuits of the TTFS imager operate asynchronously like [27]. The collected temporal information of output addresses can be used to reconstruct the scene while maintaining a wide dynamic range and performing smart functions, such as histogram equalization and scene segmentation. We will explicitly discuss this novel imager in the following sections.

3.3.1 Principle

A time-based image sensor essentially tries to extend the dynamic range on the high end, limited by the power supply for typical APSs. Now considering when the photocurrent discharges the pixel capacitance, the following relation holds

$$I_{ph} = C_{pd}\frac{\Delta V}{\Delta t_{int}} \tag{3.7}$$

where ΔV is the measured pixel output voltage or signal swing, C_{pd} is the pixel capacitance (assumed to be constant), Δt_{int} is the integration time for each pixel, and I_{ph} is the photocurrent for each pixel (assumed to be constant in each integration period). As discussed previously, typical APSs choose a fixed integration time, and the signal swing limits the typical dynamic range to about 70dB. On the contrary, the output of each pixel in time-based imagers is not a voltage but the time Δt_{int} taken to discharge the pixel capacitance. The system is no longer forced to choose a single integration time, since each pixel naturally chooses a suitable integration time. For still image mode, the dynamic range can be significantly enhanced when this temporal coding is used.

The TTFS also uses temporal coding. Instead of reading out the analog voltage across each photodiode at a predetermined exposure time, there is a comparator inside each pixel. When the voltage on a photodiode drops below a global reference voltage V_{ref}, the comparator inverts, and the pixel generates a pulse (i.e., it has "fired"). As shown in figure 3–4, three pixels are located at the addresses (*row1, col1*), (*row2, col2*) and (*row3, col3*), with the photocurrents of *I1, I2* and *I3*. Once one pixel has fired, it is disabled for the rest of the frame after its address is output. The time at which a pixel's address is read out represents the photocurrent (illuminance) of the pixel. For this scheme, an ADC, which is required for conventional imagers to output digital values, can be replaced with a simple digital counter that records the time when each pixel fires. This scheme is in fact similar to a single-slope ADC. In

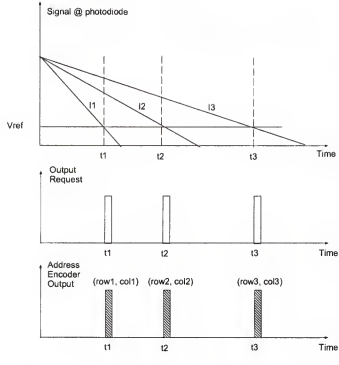


Figure 3–4: Scheme for the TTFS imager in still image mode

order to reconstruct an image, the firing time of each pixel will likely be quantized and stored in an external receiver (PC or DSP processor). This time quantization occurs after the information is scanned off the imager chip. Since the off-chip digital clock can run very fast, the amount of quantization noise is small relative to the other delays due to buffering and collisions.

Since the signal is sampled in the time domain, the dynamic range can be expressed as the ratio of the longest and shortest integration (firing) times

$$DR = 20 \log \frac{i_{max}}{i_{min}} = 20 \log \frac{t_{longest}}{t_{shortest}} \tag{3.8}$$

Usually, the dark current limits the longest possible pixel integration time (firing) time, which, from our measurement, is about $4.5s$ for the TTFS imager fabricated in TSMC $0.18\mu m$ digital technology. The shortest exposure time for a single pixel is designed to be $1\mu s$, which gives a dynamic range of more than 130dB ($20 \log(4.5s/1\mu s) =$

133dB). For a large size array, the shortest exposure time is expected to be $10\mu s$, which yields a 113dB dynamic range.

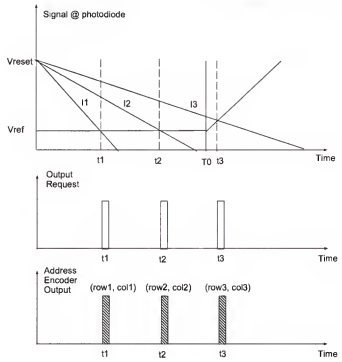


Figure 3–5: Scheme for the TTFS imager in video mode

For video mode applications, the maximum achievable integration time is limited by the video frame time, which is typically about 33ms. For this reason, current CMOS processes limit time-based (including TTFS) imagers to around 70dB of dynamic range which is about the same as for conventional CMOS imagers [34]. However, the TTFS imager can dramatically increase the DR in video mode by slowly increasing the reference voltage from the lowest value to the reset voltage during a frame. All pixels are guaranteed to pass the threshold and fire within the fixed frame time. Existing time-based imagers where the pixels are run as asynchronous oscillators are not able to take advantage of this feature.

The extended DR due to a varying voltage reference can be illustrated intuitively as follows. Taking the log of each side of equation 3.7 gives

$$\log I_{ph} = \log C_{pd} + \log \Delta V - \log \Delta t_{int} \tag{3.9}$$

The lefthand side is the pixel illuminance and its DR must be captured in the sum of the DR of the components on the righthand side. Since C_{pd} is constant, the DR can only be captured by the ΔV and Δt_{int} terms. For conventional CMOS imagers, the integration time Δt_{int} is fixed, and all of the DR must be encoded in ΔV resulting in the usual 65-75dB limitation. Typical time-based imagers hold ΔV constant and all of the DR must be encoded in Δt_{int} resulting in the 65-75dB limitation given above for video mode. For the varying V_{ref} case, the full DR is the sum of the DR of both Δt_{int} and ΔV producing a 130-150dB dynamic range for video mode, which is illustrated in figure 3–5. Obviously, the varying reference voltage favors the dynamic range extension for video mode applications. It, however, degrades the captured scene's SNR, which will be discussed later.

3.3.2 System Architecture

The original TTFS image sensor architecture is shown in figure 3–6. To differentiate it from the modified architectures, we call it the TTFS_classic imager. Inside each pixel, there is a photodiode, a comparator and a digital control circuitry. To minimize the FPN, an autozeroing technique is adopted to reset the pixel.

The TTFS_classic imager operates as follows:

1. After reset, the photocurrent discharges the photodiode. When the voltage across photodiode drops below the comparator's reference voltage, the pixel makes a row request by pulling down the $\sim row\_request(\mathrm{m})$ line.

2. The row arbiter arbitrarily selects a row from the requesting group by making $row\_select(\mathrm{m})$ high. Subsequently, this row's address is stored by the row address encoder.

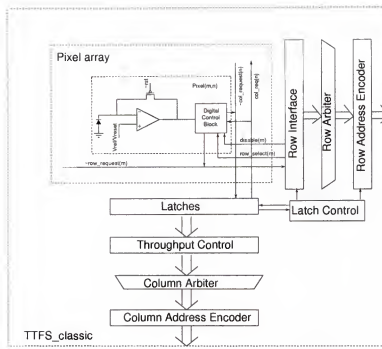


Figure 3–6: TTFS_classic imager block diagram

3. When the control signal *row_select*(m) is received, all firing pixels in that row are allowed to make column requests by pulling down the $\sim$*col_request*(n) lines.

4. The column latch records all column requests of firing pixels in the selected row.

5. Once all column requests are latched, the pixels in the selected row that had been making column requests are disabled from firing again for the rest of the frame by *disable*(m) and *col_req*(*n*) via the digital control block. As a result, $\sim$*row_request*(m) from this row is withdrawn. After that, if there are other valid $\sim$*row_request*(m) signals, new row arbitration is allowed to start taking place. However, the row interface circuit blocks a new *row_select*(m) to generate until all valid data inside latch cells have been processed. Column arbitration begins on the requests in the column latch cells. Note that the throughput control circuit is used to control the column arbitration speed.

6. When a column is selected by the column arbiter, its column address, the latched row address and the time information are all read out. The time stamp represents the illuminance information.

7. Once the column arbitration is complete, i.e., all valid data inside the latch have been processed, the row interface circuit allows a new *row_select*(m) signal to become valid. At this point, a readout cycle is finished.

3.3.3 SNR Consideration

SNR is another performance criterion, which is defined as the ratio of the input signal power and the average input referred noise power. In [35], a simplified integration mode pixel model is depicted, where both signal and noise are expressed in the form of charge. The additive noise has two independent components: (i) shot noise with zero mean and variance of $q(i_{ph}+i_{dark})t$, where t is the integration time, i_{ph} and i_{dark} are the photocurrent and dark current respectively; (ii) readout noise σ_r. Thus the SNR can be calculated as

$$SNR(i_{ph}) = \frac{(i_{ph}t)^2}{q(i_{ph}+i_{dark})t + \sigma_r^2} \tag{3.10}$$

A more complex noise model may include the FPN. For simplicity, we neglect this noise source here since the FPN can be reduced to a negligible value by using either a correlated double sampling or an autozeroing reset technique. From the above equation, we notice that SNR is also a function of integration time. Simulated SNR with respect to integration time is plotted in figure 3–7. Note that SNR monotonically increases with integration time, and ultimately reaches the upper bound due to well saturation. If not saturated, larger signals enjoy better SNR as expected.

Since each pixel in a TTFS imager can reach its full well capacity Q_{well} in still mode applications, the SNR is maximized as shown in figure 3–8. This conclusion is valid for most time-based image sensors, since they all equivalently use a constant reference voltage scheme.

In video-mode TTFS image sensors, pixels with weak photocurrents cannot integrate for too long to reach their well capacities. Now consider the case that the varying reference voltage is piece-wise linear as shown in figure 3–5. A pixel with a strong photocurrent such that the firing time is less than T_0 can achieve full well capacity, whereas a pixel with a weaker photocurrent, for instance the pixel (*row3, col3*) in figure 3–5, does not. If ignoring dark current, the firing time for a weak

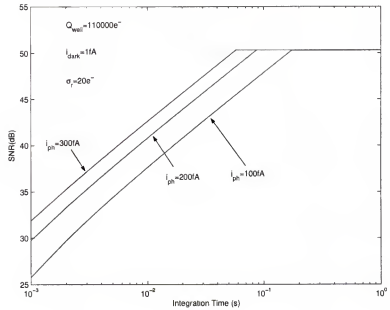


Figure 3–7: SNR vs. integration time

signal, e.g., $i_{ph} < Q_{well}/T_0$, is

$$t = \frac{(Q_{well} + kT_0)}{i_{ph} + k} \tag{3.11}$$

where k is the slope of the linearly increasing part of the reference voltage, $k = \dfrac{Q_{well}}{t_{frame} - T_0}$. Then the specific expression of SNR is given as following

$$SNR(i_{ph}) = \begin{cases} \dfrac{\left(i_{ph}\dfrac{Q_{well} + kT_0}{i_{ph} + k}\right)^2}{i_{ph}\dfrac{Q_{well} + kT_0}{i_{ph} + k}q + \sigma_r^2} & i_d \le i_{ph} \le \frac{Q_{well}}{T_0} \\ \dfrac{Q_{well}^2}{qQ_{well} + \sigma_r^2} & i_{ph} \ge \frac{Q_{well}}{T_0} \end{cases} \tag{3.12}$$

Figure 3–8 shows the SNR comparison of TTFS imagers with different reference voltage schemes. The comparison was made by assuming a well capacity of 110,000 electrons, a readout noise of 20 electrons and a dark current of 1fA. It indicates that the imagers with constant reference voltage enjoy better averaged SNR, since weak signals have taken advantage of full well capacity. Even though the varying

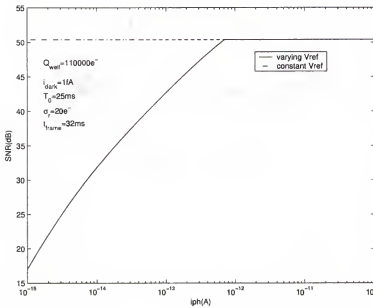


Figure 3–8: SNR of TTFS imagers

reference voltage scheme favors dynamic range extension for video mode applications, it degrades the captured scene's average SNR, especially for the dark region. It turns out that T_0 in figure 3–5 is a key factor affecting the SNR performance. A simple optimal strategy for piece-wise linear reference voltage variation will be presented in the next chapter.

For video mode applications, we investigate the SNRs for the multiple sampling technique and the time-based method. The simulation for multiple sampling was done for a sensor with well capacity of 110,000 electrons, readout noise of 20 electrons, dark current of 1fA and 9 captures at $32ms/2^8$, $32ms/2^7$, $\cdots$, $32ms$. The simulation of the time-based imager is performed for a sensor with the same well capacity, readout noise and dark current, except for a frame time of 32ms and a reference voltage change point at 25ms. The simulation results are given in figure 3–9. We observe that the time-based imager benefits the SNR in the strong signal region. The SNR of the multiple sampling technique shows numerous 3dB dips since each pixel cannot take advantage of the full well capacity [35]. It should be pointed out here, for a

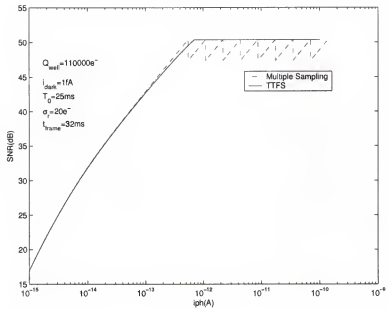


Figure 3–9: SNR comparison of multiple sampling imagers and TTFS imagers

multiple sampling imager with APS architecture, the SNR is about 2-3dB worse than the simulated results due to the narrowed signal swing as discussed in [36].

3.4 Summary and Discussion

The dynamic range limitation of conventional CMOS APSs has been discussed in this chapter. Several existing dynamic range enhancement methods were reviewed. A novel time-to-first-spike CMOS imager was introduced in detail, which has demonstrated superior performances, e.g., DR and SNR, compared to conventional CMOS APSs. In addition, we summarize the comparison between the TTFS_classic imager and existing high dynamic enhancement methods in table 3–1. One unique issue with the TTFS_classic imager is the situation when many pixels send out *request* signals within a short period. The asynchronous readout circuit has to deal with this collision problem, and some amount of delay will be inevitably introduced resulting in some temporal errors for the reconstructed illumination. We will address this issue and propose some architectures to solve this problem in Chapter 6.

Table 3–1: Comparison of the TTFS_classic imager and existing high dynamic range enhancement methods

Sensor Type	SNR	Power	Memory	Video	Collision
NSC	Low	Low	Few	Yes	No
MS	Good	High	More	Yes	No
TBAM	Good	High	More	No	No
TBSM	Good	High	Few/More	Yes	No
TTFS_classic	Good	Low	Few	Yes	Yes

The previous discussed dynamic range enhancement methods all attempt to extend the dynamic range towards the strong signal region. As the other choice of DR enhancement, the extension over the weak signal region can be achieved by either a dark current reduction approach [15] or a statistic signal processing method [37].

CHAPTER 4
AN OPTIMAL TWO SEGMENT PIECE-WISE LINEAR STRATEGY FOR REFERENCE VOLTAGE VARIATION

For video mode applications, the maximum integration time is limited by the frame time, which is typically 33ms. Dynamic range would be limited if using a constant reference voltage, as discussed in the previous chapter. By varying the reference voltage, all the useful illuminance information can be collected within each frame period. So far, there has not been any systematic study of what is the optimal strategy for the reference voltage variation. T. Chen [38] has studied an optimal scheduling of capture times for the multiple sampling method by assuming the complete incident illumination probability density function (pdf) is known in advance. In [31], a sampling strategy has been discussed for a time-based imager with synchronized readout scheme by considering the shortest sampling interval. These two papers both try to acquire an optimal solution with the objective of achieving maximum average SNR under some constraints. For a TTFS imager, the reference voltage is running in the continuous mode, therefore we cannot set individual capture times as required in the above two cases. In this chapter, we will systemically analyze an optimal two segment piece-wise linear strategy for the reference voltage (or equivalently, well capacity) variation using the expected SNR as the objective. In the following sections, we will first formulate this optimization problem. An *off-line* optimal strategy is discussed in section 4.2, provided the required minimum time interval and the real comparator-introduced time delay are given. In section 4.3, an *on-line* optimal strategy is given by incorporating the nonuniform quantization noise.

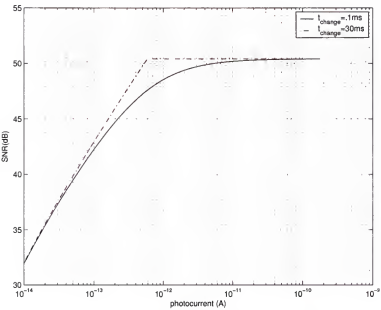


Figure 4–1: SNR vs. photocurrent for different t_{change}s. Here, $Q_{well} = 110,000e^-$ and $t_{frame} = 33ms$.

4.1 Problem Formulation

Including all noise sources, SNR can be expressed as

$$SNR(i_{ph}) = \frac{(i_{ph}t)^2}{(i_{ph} + i_{dark})tq + \sigma_r^2 + \sigma_Q^2 + \sigma_{FPN}^2} \tag{4.1}$$

where σ_Q and σ_{FPN} are the quantization noise and the FPN respectively. It has been shown earlier that a TTFS imager enjoys better SNR than conventional CMOS APS imagers. However, this achievement depends on reference voltage selection (see figure 4–1). Thus, our optimization problem can be formulated as follows:

For a two segment piece-wise linear reference voltage, find the optimal t_{change}, which maximizes the expected SNR under some constraints.

In the following discussion, we will address well capacity instead of reference voltage (see figure 4–2).[1]

[1] t_{change} in figure 4–2 is interchangeable with T_0 in figure 3–5.

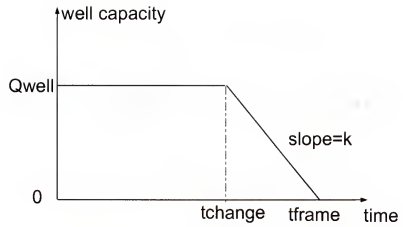


Figure 4–2: Varying well capacity for a wide dynamic range

4.2 Optimal Strategy for Uniform Quantization

Suppose we uniformly quantize the incident light intensity, which indicates that the quantization noise σ_Q^2 is the same for different photocurrents. In addition, for simplicity of discussion, we ignore all the noise terms except for the photocurrent shot noise. If we define $i_{change} = Q_{well}/t_{change}$, then SNR as a function of photocurrent is

$$SNR(i_{ph}) = \begin{cases} \dfrac{Q_{well}}{q} & \text{if } i_{ph} \geq i_{change} \\ \dfrac{i_{ph}(Q_{well} + kt_{change})}{q(i_{ph} + k)} & \text{if } i_{ph} < i_{change} \end{cases}$$

Note that a large photocurrent's SNR reaches the maximum value, while a small photocurrent's SNR directly depends on t_{change}, or the slope $k = \dfrac{Q_{well}}{t_{frame} - t_{change}}$.

4.2.1 Optimal Strategy under Minimum Time Interval Constraint

If we suppose the incident illumination pdf $f_I(i)$ is zero outside of a finite length interval (i_{min}, i_{max}), where $i_{min} < Q_{well}/t_{frame}$ and $i_{max} > Q_{well}/t_{frame}$, then this optimization problem can be expressed as follows:

For a two segment piece-wise linear well capacity, given the photocurrent quantization step ΔI and the required minimum time interval Δt_{min}, find the optimal

t_{change}, which maximizes the expected SNR

$$E(SNR) = \int_{i_{change}}^{i_{max}} \frac{Q_{well}}{q} f_I(i)\, di + \int_{i_{min}}^{i_{change}} \frac{i(Q_{well} + kt_{change})}{q(i+k)} f_I(i)\, di \qquad (4.2)$$

subject to $i_{min} \leq i_{change} \leq i_{max}$, and any time interval Δt between two adjacent quantized photocurrents' firing times is no less than Δt_{min}, i.e., $\Delta t \geq \Delta t_{min}$.

The minimum time interval Δt_{min} is introduced by the hardware implementation limitation. Any two adjacent quantized photocurrents' firing times should widely spread, otherwise we cannot separate them apart in the time domain. Obviously for all photocurrents, $SNR(i_{ph}) \leq Q_{well}/q$. Thus the upper bound of $E(SNR)$ is Q_{well}/q. If $i_{min} \geq Q_{well}/t_{frame}$, then the optimal solution is simply $i_{change} = Q_{well}/t_{frame}$, and

$$E(SNR) = \int_{i_{min}}^{i_{max}} \frac{Q_{well}}{q} f_I(i)\, di = \frac{Q_{well}}{q} \qquad (4.3)$$

For this case, the expected SNR achieves its upper bound. Actually, such scene information about the incident light intensity should be known in advance, so it is generally an *on-line* method. However, our goal here is to provide an *off-line* strategy. Suppose i_{min} is less than i_{change}, and we rearrange the objective function as follows:

$$\begin{aligned} E(SNR(i_{change})) &= \frac{Q_{well}}{q} - \int_{i_{min}}^{i_{change}} \frac{k(Q_{well} - it_{change})}{q(i+k)} f_I(i)\, di \\ &= \frac{Q_{well}}{q} - \int_{i_{min}}^{i_{change}} \frac{Q_{well}(1 - i/i_{change})}{q(i/k+1)} f_I(i)\, di \end{aligned} \qquad (4.4)$$

The above equation shows that if we could make k as large as possible and simultaneously make i_{change} as small as possible, the expected SNR could be maximized even though the exact information of pdf $f_I(i)$ is unknown (see figure 4–3). As we discussed earlier, the beauty of a still mode time-based image sensor is to allow each pixel to reach the full capacity, thus resulting in a high DR and good SNR. So for video mode applications, we also need to make each pixel's available well capacity as large as possible. Figure 4–4 shows that the expected SNR monotonically increases

with t_{change} by assuming the scene has a uniform pdf. Even though the scene pdf will not be uniform in practice, a large t_{change} (or slope k) can still guarantee that each pixel will reach its maximum well capacity and achieve the best SNR.

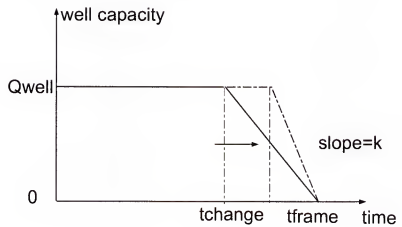


Figure 4–3: Illustration of optimization

The general solution derived in the previous discussion indicates that the optimal solution for slope k is infinity without any other constraints. However, in this case, all the photocurrents below i_{change} will be quantized to an undesirable single value. To solve this problem, we introduce the *minimum required time interval constraint* here, named Δt_{min}. Two photocurrents with ΔI difference can be differentiated in the time domain as long as their firing times' interval is larger than Δt_{min}. Suppose we have two photocurrents I_m and I_{m+1}, which are related by

$$I_{m+1} = I_m + \Delta I \tag{4.5}$$

Then if the well capacity stays constant, we find that the time interval Δt_m between the two photocurrents is

$$\Delta t_m \equiv t_m - t_{m+1} = \frac{Q_{well} \cdot \Delta I}{I_{m+1} I_m} \tag{4.6}$$

Clearly, a large photocurrent I_m implies a small time interval Δt_m. In order to obtain all the useful information, we must ensure each time interval is larger than the lower

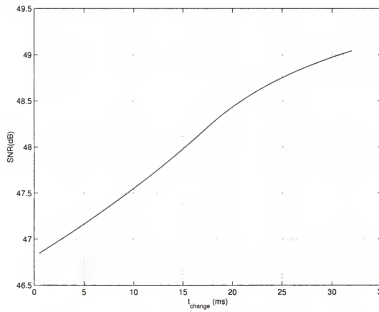


Figure 4–4: Expected SNR vs. t_{change}. Here, $Q_{well} = 110,000e^-$ and $t_{frame} = 33ms$. Assume uniform pdf $f_I(i), i \in (10fA, 1000fA)$.

bound Δt_{min}. A detailed analysis of the time intervals of large photocurrents can be found in [31]. Here we simply skip this part. For the photocurrent pair I_m and I_{m+1} that are both less than i_{change}, their time interval is

$$\Delta t_m \equiv t_m - t_{m+1} = \frac{(Q_{well} + kt_{change}) \cdot \Delta I}{(I_{m+1} + k)(I_m + k)} \tag{4.7}$$

It can be found that the shortest time interval is around the point where the well capacity starts to change. Now considering the photocurrent pair i_{change} and $i_{change} - \Delta I$, we have

$$\Delta t_{worst} = \frac{Q_{well} + kt_{change}}{i_{change} - \Delta I + k} - \frac{Q_{well}}{i_{change}} \tag{4.8}$$

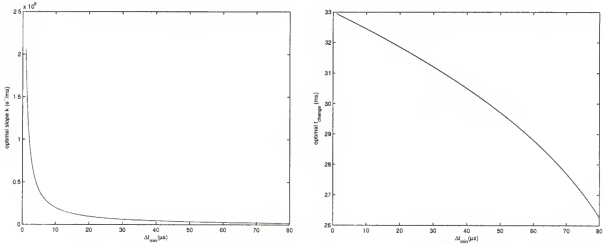
In general, $k \gg \Delta I$, then it yields

$$\Delta t_{worst} = \frac{\Delta I(kt_{frame} - Q_{well})^2}{k^3 t_{frame}} \tag{4.9}$$

Let $\Delta t_{worst} > \Delta t_{min}$, then we have the following inequality

$$\frac{\Delta I(kt_{frame} - Q_{well})^2}{k^3 t_{frame}} > \Delta t_{min} \tag{4.10}$$

It provides an upper bound k_{max1} for the slope k with respect to the required minimum time interval Δt_{min} and the given photocurrent resolution ΔI. In other words, k_{max1} is the optimal solution. Intuitively, a small required minimum time interval leads to a large optimal slope k, which means an infinitely fast clock and an infinitely large throughput result in an infinite k. Figure 4–5 verifies this intuition. In contrast, figure 4–6 shows that the optimal solution monotonically decreases with the photocurrent resolution ΔI. That is, we need to extend the time interval for two adjacent photocurrents with small difference.

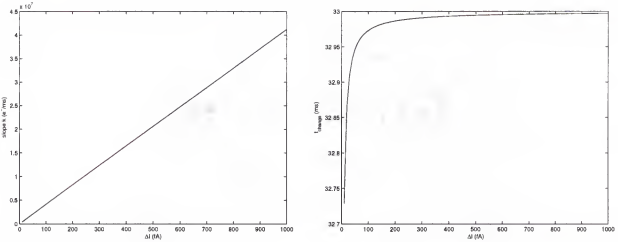


(a) (b)

Figure 4–5: Effect of Δt_{min} on the optimal solution for uniform quantization. (a) effect of Δt_{min} on k_{max1}. (b) effect of Δt_{min} on t_{change}. Here, $t_{frame} = 33ms$, $\Delta I = 10fA$ and $Q_{well} = 110,000e^-$.

4.2.2 Comparator Delay Considerations

In a typical TTFS imager design, the pixel level comparator also works as an opamp to implement an autozeroing technique for the offset FPN reduction. Also due to power conservation considerations, the resulting -3dB frequency is so low



(a) (b)

Figure 4–6: Effect of ΔI on the optimal solution for uniform quantization. (a) effect of ΔI on k_{max1}. (b) effect of ΔI on t_{change}. Here, $t_{frame} = 33ms$, $\Delta I = 10fA$ and $Q_{well} = 110,000e^{-}$.

that comparator-introduced time delay is not negligible. So in this section, we will investigate the effect of comparator-introduced delay on firing times.

According to the small-signal model of comparator, when the input signal approaches the reference signal, a comparator can be viewed as a low-pass filter given by

$$f(s) = \frac{A}{1 + s/w_p} \tag{4.11}$$

where A is the small-signal comparator gain, and w_p is the -3dB frequency. Assuming a photocurrent i, a photodiode associated capacitance C and a small signal region $|v_{in}| \leq a$, then the small signal of the comparator input can be seen as

$$v_{in}(t) = h \cdot t - a \tag{4.12}$$

where $h = i/C$. According to Kirchhoff's current law, this system can be also described as a differential equation

$$\frac{dv_{out}(t)}{dt} + \frac{v_{out}(t)}{\tau} = \frac{Av_{in}(t)}{\tau} \tag{4.13}$$

Here $\tau = 1/w_p$, and $v_{out}(\cdot)$ is the small output signal. The solution to this well-known differential equation turns out to be

$$v_{out}(t) = Av_{in}(t) - Ah\tau + ce^{-t/\tau} \tag{4.14}$$

With the initial conditions $v_{out}(0) = -b, v_{in}(0) = -a$ and $b \approx Aa$, we have $c = Ah\tau + Aa - b \approx Ah\tau$. Then the resulting $v_{out}(t)$ will be

$$v_{out}(t) = Av_{in}(t) - Ah\tau(1 - e^{-t/\tau}) \tag{4.15}$$

If we define t_1 as the time when the input reaches the reference voltage (i.e., the small signal $v_{in}(t_1) = 0$) and suppose the output small signal reaches 0 at t_2, we have

$$\Delta t_d \equiv t_2 - t_1 = \tau \cdot (1 - e^{-t_2/\tau}) \tag{4.16}$$

where Δt_d is the time delay between t_2 and t_1. The key observation here is that the time delay will roughly equal τ if $t_2 \geq 4\tau$. This has been verified by our CADENCE simulations. The analytic solution of t_2 can be obtained from the following

$$ht_2 - a = h\tau(1 - e^{-t_2/\tau}) \tag{4.17}$$

It shows that t_2 is a function of h or i, so is Δt_d. From the time delay point of view, an increased i will result in a small delay. Provided that the real comparator delay is on the same order of the required minimum time interval Δt_{min}, we have to include this *time delay constraint* under consideration. The real readout time stamp now becomes $T' = T + \Delta t_d(i)$, where T is the ideal firing time, and $\Delta t_d(\cdot)$ is the real comparator-introduced delay. As before, the serious effect of the comparator delay on the time interval happens around the varying well capacity start point. To see this effect, we rewrite the worst time interval as

$$\Delta t'_{worst} = \frac{\Delta I(kt_{frame} - Q_{well})^2}{k^3 t_{frame}} + \Delta t_d(i_{change} - \Delta I) - \Delta t_d(i_{change}) \tag{4.18}$$

Assume $\Delta t_d(i_{change})$ is only decided by i_{change}, while both $i_{change} - \Delta I$ and slope k are the contributors for $\Delta t_d(i_{change} - \Delta I)$. As k increases, $\Delta t_d(i_{change} - \Delta I)$ becomes much less than $\Delta t_d(i_{change})$, which in turn results in a much smaller actual time interval. In this case, inequality 4.10 no longer holds. To avoid this problem, there should be another upper bound for slope k, which ensures the comparator-introduced time delay for each photocurrent is roughly the same. As shown previously, if t_2 is larger than 4τ, then the time delay can be treated as a constant τ. Now replacing h with $(i_{change} - \Delta I + k)/C$, equation 4.17 yields

$$(i_{change} - \Delta I + k)t_2/C - a = (i_{change} - \Delta I + k)/C \cdot \tau(1 - e^{-t_2/\tau}) \tag{4.19}$$

To ensure t_2 greater than 4τ, it must have[2]

$$i_{change} - \Delta I + k \leq \frac{a \cdot C}{3\tau} \tag{4.20}$$

Substituting i_{change} with $\dfrac{Q_{well}}{t_{frame} - Q_{well}/k}$ and assuming $i_{change} \gg \Delta I$, then the above inequality becomes

$$k^2 - \frac{a}{3\tau}C \cdot k + \frac{Q_{well}}{t_{frame}}\frac{a}{3\tau} \cdot C \leq 0 \tag{4.21}$$

It gives us another upper bound k_{max2}, which is

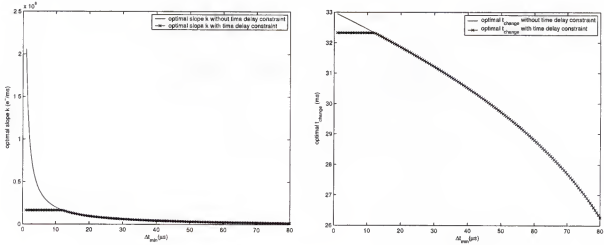
$$k_{max2} = \frac{1}{2}\left(\frac{a}{3\tau}C + \sqrt{\frac{a}{3\tau}C \cdot \left(\frac{a}{3\tau}C - 4\frac{Q_{well}}{t_{frame}}\right)}\right) \tag{4.22}$$

The two upper bounds for the slope k occur under two different constraints. We simply take the minimum of the two upper bounds as our final solution for this optimization problem, which is

$$k_{opt} = \min(k_{max1},\ k_{max2}) \tag{4.23}$$

[2] In the derivation, we make the following approximation $1 - e^{-4} \approx 1$.

From the simulation results shown in figure 4–7, we observe that the optimal solution is merely decided by the required minimum time interval when it is large enough. Meanwhile, the time delay constraint plays a key role in determining the optimal solution if the required minimum time interval is much smaller. In figure 4–8, optimal solution vs. ΔI is plotted. It shows that the required minimum time interval constraint dominates in the small ΔI region, whereas the optimal solution is mainly decided by the time delay constraint for a large ΔI.



(a) (b)

Figure 4–7: Optimal slope k and t_{change} vs. Δt_{min} under time delay constraint. Here, $a = 10mV$, $t_{frame} = 33ms$, $\tau = 1\mu s$, $\Delta I = 10fA$ and $Q_{well} = 110,000e^-$.

4.2.3 Optimization Steps for Uniform Quantization

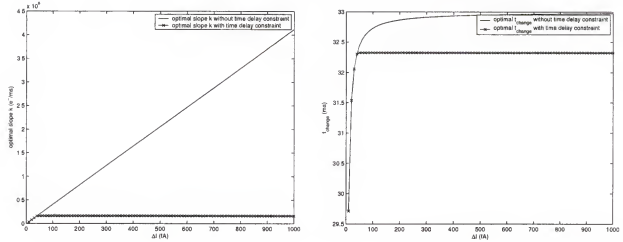
In summary, this optimization problem can be solved with the following steps:

1. Compute the upper bound k_{max1} under the required minimum time interval constraint for the ideal comparator case using inequality 4.10;

2. Compute the upper bound k_{max2} under the time delay constraint using inequality 4.22;

3. Take the minimum of the two upper bounds as the final optimal solution k_{opt}, and the optimal t_{change} is related by

$$t_{change} = t_{frame} - \frac{Q_{well}}{k_{opt}}$$



(a) (b)

Figure 4–8: Optimal slope k and t_{change} vs. ΔI under time delay constraint. Here, $a = 10mV$, $t_{frame} = 33ms$, $\tau = 1\mu s$, $\Delta t = 50\mu s$ and $Q_{well} = 110,000e^{-}$.

4.2.4 Discussion

This optimization method does not need any specific information about the scene illumination, so it is an *off-line* method. Certain scene illumination information, however, is helpful to solve this optimization problem. When the readout noise and the FPN are not negligible, the upper bound of the expected SNR will decrease, but the final optimal solution still holds. The constraints, we consider here, are all time related. In addition, if a high-speed comparator with a time constant of several ns is used, the time delay will no longer be an issue. To achieve a large safety margin, we can let $t_2 \geq 5\tau$. Then the upper bound k_{max2} will be

$$k_{max2} = \frac{1}{2}\left(\frac{a}{4\tau}C + \sqrt{\frac{a}{4\tau}C \cdot \left(\frac{a}{4\tau}C - 4\frac{Q_{well}}{t_{frame}}\right)}\right) \tag{4.24}$$

4.3 Optimal Strategy for Nonuniform Quantization

Pixel-parallel automatic gain control is an inherent characteristic of time-based image sensors [27]. Uniform quantization along the time axis will result in *nonuniform* quantization in the photocurrent domain. Generally, nonuniform quantization is

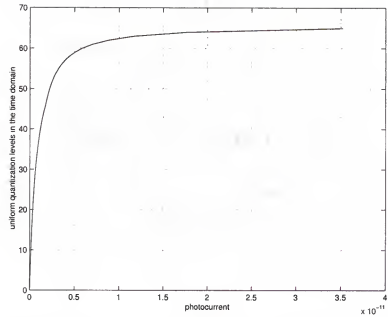


Figure 4–9: Mapping between photocurrents and firing times. Here, $\Delta_{time} = 0.5ms$, $t_{change} = 10ms$ and $t_{frame} = 33ms$

beneficial to reduce the quantization noise for the predominantly weak photocurrents at the expense of an increase in noise for the rarely occurring strong photocurrents. The relationship between photocurrents and corresponding quantization values in the time domain is illustrated in figure 4–9. It has a typical compression characteristic, having a much steeper slope for small magnitude photocurrents than that for large magnitude photocurrents. Thus a given signal change in the small magnitude region will carry the uniform quantizer through more steps than the same change in the large magnitude region. Similar to standard compression curves used in voice transmission, e.g., μ-Law and A-Law, varying the reference voltage also aims to achieve continuity at the origin of the photocurrent. Unlike the uniform quantizer, it is not suitable for nonuniform quantizer to use a single average quantization noise power for different input signal levels. For a midrise quantizer, the mean squared error (MSE) for a signal in the r^{th} quantizing interval is computed as

$$E_q(r) = \int_{x_r - \Delta_r/2}^{x_r + \Delta_r/2} (x - x_r)^2 p_r \, dx \tag{4.25}$$

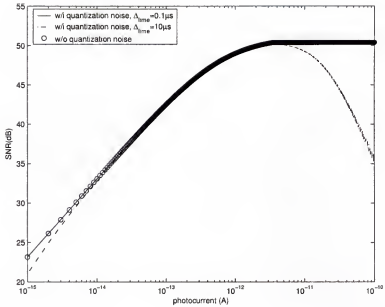


Figure 4–10: SNR vs. photocurrent when including the quantization noise. Here, $t_{frame} = 33ms$, $t_{change} = 5ms$ and $Q_{well} = 110,000e^{-}$.

where Δ_r is the step size of r^{th} quantizing interval, and p_r is the input signal pdf. Assuming the signal is uniformly distributed over $[x_r - \Delta_r/2,\ x_r + \Delta_r/2]$, we have

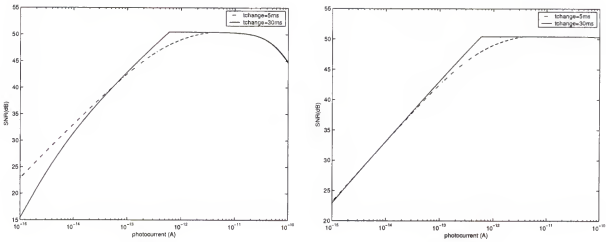
$$E_q(r) = \frac{\Delta_r^2}{12} \tag{4.26}$$

Here we regard MSE as the quantization noise, which is

$$\begin{gathered} \sigma_Q^2(i) = \frac{\Delta_r^2}{12} \\ \Delta_r = I_{r+1} - I_r \text{ and } I_r \leq i < I_{r+1} \end{gathered} \tag{4.27}$$

where I_r and I_{r+1} are quantization levels, decided by the well capacity varying curve and the quantization step size Δ_{time} in the time domain. They can be expressed as

$$I_r = \begin{cases} \dfrac{Q_{well}}{r \cdot \Delta_{time}} & \text{if } r \cdot \Delta_{time} \leq t_{change} \\ \dfrac{Q_{well}(t_{frame} - r \cdot \Delta_{time})}{(t_{frame} - t_{change})(r \cdot \Delta_{time})} & \text{if } r \cdot \Delta_{time} > t_{change} \end{cases} \tag{4.28}$$



(a) (b)

Figure 4–11: Comparison of SNRs for different t_{change}s. (a) $\Delta_{time} = 3\mu s$, (b) $\Delta_{time} = 0.3\mu s$. Here, $t_{frame} = 33ms$ and $Q_{well} = 110,000e^{-}$.

The above equation indicates that the well capacity varying curve not only decides the quantization levels but also determines the quantization noise. Keeping the quantization noise and the photocurrent shot noise, we rewrite the expression of SNR as follows:

$$SNR(i) = \frac{(it)^2}{itq + \sigma_Q^2(i) \cdot t^2} \tag{4.29}$$

The new SNR expression is plotted in figure 4–10. The simulated SNR illustrates that when the quantization noise is quite small, the shot noise dominates the noise power, and when the quantization noise is comparable to the shot noise, SNR is dramatically degraded. If a long t_{change} is applied, a large quantization step size is achieved for the small magnitude photocurrents, which also leads to a large quantization noise and a worse SNR. On the contrary, the same long t_{change} benefits the SNR of the medium magnitude photocurrents thanks to the large achievable well capacity (see figure 4–11 (a)).

Including the quantization noise, our optimization problem becomes

For a two segment piece-wise linear well capacity, given the time quantization step size Δ_{time} and the incident light pdf $f_I(i)$, find the optimal t_{change}, which maximizes

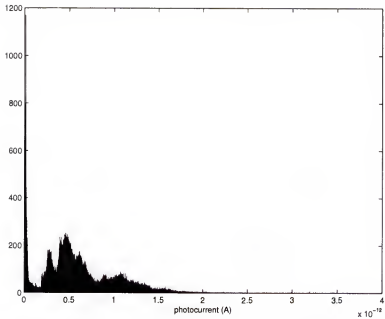


Figure 4–12: Photocurrent histogram of *vinesunset*

the expected SNR

$$E(SNR) = \int_{i_{min}}^{i_{max}} SNR(i) \cdot f_I(i)\, di \tag{4.30}$$

subject to $0 \leq t_{change} \leq t_{frame}$.

where $SNR(i)$ is given in equation 4.29. If Δ_{time} is small enough, $SNR(i)$ will reduce to $\frac{it}{q}$ (see figure 4–10 and figure 4–11(b)). The optimization here needs to know the pdf of the incident light, so it is in general an *on-line* method. This rough scene information may be gathered during the previous captures. In the following discussion, we will investigate the optimization problem by assuming the complete scene illuminance information is known in advance.

Suppose $f_I(i)$ is zero outside $(i_{min},\ i_{max})$. As before, if $i_{min} \geq Q_{well}/t_{frame}$, then the solution of equation 4.30 is simply $t_{change} = t_{frame}$. A decreased t_{change} increases the small signal's SNR but degrades the large signal's SNR. Besides, the expected SNR (ESNR) depends on not only t_{change} but also the incident intensity pdf. Though few natural scenes exhibit uniform illumination statistics, any pdf can

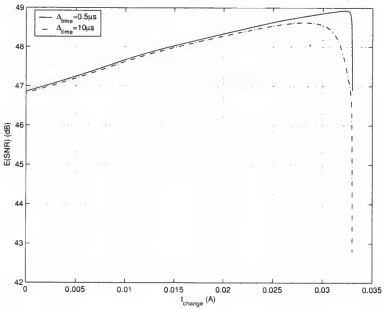


Figure 4–13: ESNR vs. t_{change}. For $\Delta_{time} = 0.5\mu s$, optimal $t_{change} = 32.3ms$; and for $\Delta_{time} = 10\mu s$, optimal $t_{change} = 28.2ms$.

be approximated by a piece-wise uniform pdf. Then assuming that the pdf is uniform over $(i_{min1}, i_{max1}), \cdots, (i_{minN}, i_{maxN})$, the objective function becomes:

$$E(SNR) = \sum_{k=1}^{N} p_k \cdot \int_{i_{mink}}^{i_{maxk}} \frac{(it)^2}{itq + \sigma_Q^2(i) \cdot t^2}\, di \tag{4.31}$$

where the constant p_k is the pdf value over the range (i_{mink}, i_{maxk}).

Vinesunset is a high dynamic range image, which will be introduced in Chapter 6. Its original pixel values are expressed with floating point values and proportional to the true light intensities. In our simulation, each value is simply scaled by $100\times i_d$, where i_d is the dark current, i.e., $i_d = 1fA$. Figure 4–12 shows the histogram of the scaled *vinesunset's* light intensities, which can be approximated by a five segment piece-wise uniform pdf, i.e., $(i_d,\ 40i_d)$, $(40i_d,\ 200i_d)$, $(200i_d,\ 750i_d)$, $(750i_d,\ 1500i_d)$ and $(1500i_d,\ 1900i_d)$. We plotted equation 4.31 with respect to t_{change} in figure 4–13, which is obviously a convex curve. The optimal solution is the t_{change} corresponding to the maximum ESNR. Two different Δ_{time}s are compared here, of which the smaller

step size enjoys the better ESNR because of its lower quantization noise. Additionally, figure 4–13 also shows that with integration time increasing, an enlarged well capacity benefits ESNR as expected, and when t_{change} becomes too large, quantization noise is pronounced and ESNR dramatically drops.

4.4 Summary

This chapter presented a systematic study of an optimal two segment piece-wise linear strategy for the reference voltage (or well capacity) variation. For uniform quantization, an optimal solution is achieved by considering the minimum time interval requirement and the real comparator delay. This *off-line* method does not require any explicit incident light intensity information. A time-based method essentially implements nonuniform quantization of photocurrent, which benefits weak signals' SNR, and therefore the expected SNR. Given the pdf of photocurrents, the optimal t_{change} can be determined by maximizing the ESNR. An approximate scene statistics could be achieved by the previous captures. When the reset noise and the FPN are not negligible, the ESNR will be further degraded, however the discussed optimal strategy still holds.

CHAPTER 5
TTFS_CLASSIC IMAGER DESIGN

A 128×128 TTFS_classic imager was designed using TSMC $0.18\mu m$ digital CMOS technology. In this chapter, the design is discussed at the circuit level. Section 5.1 details the pixel design, followed by single pixel test results presented in section 5.2. A detailed description of the asynchronous readout circuitry design and layout considerations are included in section 5.3 and section 5.4 respectively. Section 5.5 demonstrates the prototype chip test results. Finally, this chapter concludes with section 5.6.

5.1 Pixel Design

5.1.1 Pixel Operation

The pixel schematic of the TTFS_classic imager is shown in figure 5–1, which contains a photodiode, a comparator and a digital control circuit. Note that the signals labelled with $\sim$ are active-low. A global control logic provides V_{ref} or V_{reset} to each pixel according to the control signal $\sim rst$. There are two phases in each frame, i.e., reset phase and comparison phase. The pixel works as follows:

1. The photodiode is initially reset to V_{reset} via a negative feedback loop.

2. After the $\sim rst$ goes high, the photodiode is discharged and the voltage across the photodiode linearly drops. When the voltage drops below V_{ref}, the comparator output node flips and the node req goes high. Then the pixel sends out a request to the row arbiter by pulling down $\sim row\_request(m)$.

3. If the row arbiter selects this row by making $row\_select(m)$ high, column request $\sim col\_request(n)$ signals will be sent to the column latch.

4. After $\sim col\_request(n)$ signals are latched, the corresponding control signals, $disable$(m) and $col\_req(n)$, are generated to disable the pixel by switching on

transistor *M2*. Thus the pixel will not fire again until the next reset phase turns off *M2*.

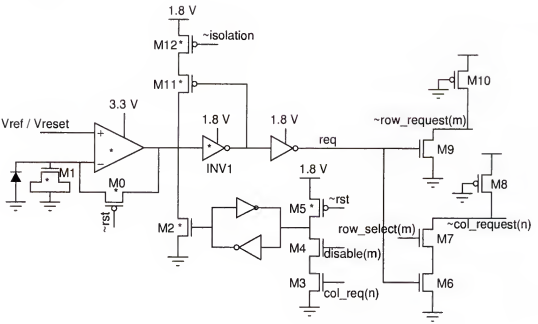


Figure 5–1: Pixel schematic of the TTFS_classic imager

We implement the front-end circuit using standard thick oxide (3.3V) transistors (labelled with *) to avoid the high gate and subthreshold leakage currents. Implementing the comparator using thick oxide transistors also makes it possible to use the high power supply (3.3V) to increase the signal swing. In order to shift down the high voltage to nominal 1.8V supply, $INV1$ with thick oxide transistors is included working as a level shifter. The positive feedback following the comparator is used to make the comparator output node immune to the switching noise. Note that the control signal $\sim isolation$, which is the same as rst except for having a delayed falling edge, is used to disable the positive feedback during the reset phase. The pixel layout is shown in figure 5–2, which is $12.4 \times 12.1 \mu m^2$ with $2.99 \times 2.97 \mu m^2$ light sensing area and has a fill factor of about 6%. Pixels are mirrored in the array in order to share the n-well and some of power and bias lines. To prevent light-induced currents from affecting the analog circuitry or causing latchup, we used the sixth layer metal everywhere except over the photosensitive node.

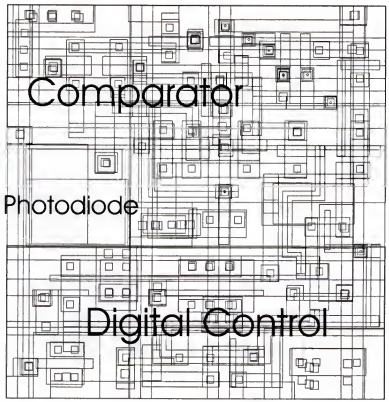


Figure 5–2: Pixel layout of the TTFS_classic

Our first pixel design in TSMC 0.18μm digital CMOS technology contains a slow opamp/comparator without positive feedback, as shown in figure 5–3. The testing results show that the firing time is inversely proportional to the light intensity. However, there exists a problem. Several unexpected oscillations or pulses come out following the expected one. It turns out that the period of oscillation is related to the light intensity, that is the weaker the light, the longer the oscillation. Including the corner analysis, we re-simulate the pixel with CADENCE, and found that the oscillation is due to the voltage of the comparator output node, especially for the Slow NMOS and Slow PMOS corner. The reason is that when the comparator flips, in some period the output voltage happens to be in the transition region of the following inverter. Since the PADOUT pad has parasitic inductance, when the huge transient DC current generates, the chip GND or VDD will oscillate, and that in turn will cause the output node of the comparator to oscillate. This is the so called Simultaneous

Switching Noise (SSN) impact. Since a weaker light generates a smaller current, the diode discharges slowly, therefore the comparator output stays in the transition region longer. To solve this problem, we adopted a relative fast opamp/comparator followed by a positive feedback in our later design, which will be analyzed in section 5.1.4. Both the fast comparator and the positive feedback are helpful to speed up the comparator out voltage transition, thus reducing the oscillations.

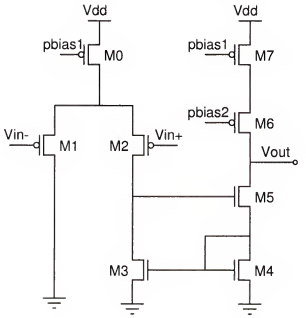


Figure 5–3: 8-transistor opamp

5.1.2 Photodiode Design

As discussed in Chapter 2, the photodiode can be formed by psub/nwell, psub/n+ or p+/nwell. To widen the depletion region to favor the photocurrent generation, we choose a psub/nwell diode shown in figure 5–4 for our TTFS_classic imager.

The total capacitance at the cathode of the photodiode is the summation of the junction capacitance of the psub/nwell diode, the comparator input PMOS transistor's gate capacitance, the MOS capacitor C_{M1}, and the drain capacitance of the

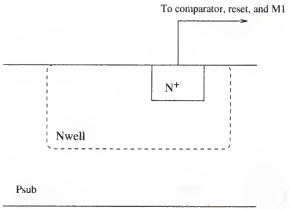


Figure 5–4: Diagram of a psub/nwell photodiode

reset PMOS transistor $M0$, which is

$$C_{pd} = C_j + C_g + C_{M0,drain} + C_{M1} \tag{5.1}$$

According to the MOSIS provided electrical parameters of the TSMC $0.18\mu m$ digital CMOS technology, i.e., the unit area capacitance between nwell and substrate is $70aF/\mu m^2$, C_j is approximately $0.62fF$, which is quite small compared to other components in C_{pd}. The MOS capacitor C_{M1} may exhibit strong voltage modulation effects [39]. Figure 5–5 shows the simulated capacitance curve for C_{M1} when the gate bias is positive. We observe that beyond strong inversion, i.e., $V_{bias} > 1$V, the capacitance is nearly constant. As discussed in Chapter 2, to achieve good linearity for the photodiode output voltage with respect to the integration time, the MOS capacitor C_{M1} has to be constant. Then this bias dependent property of the MOS capacitor sets the lower bound of reference voltage, i.e., $V_{ref} = 1$V. In addition, an increased well capacity has two other advantages: first, increase SNR and second, push strong light intensities away from the short firing time region, thus obtaining lower reconstruction errors.

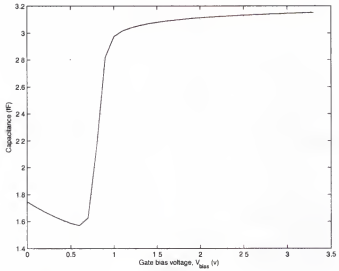


Figure 5–5: General behavior of the MOS capacitor C_{M1}

5.1.3 Autozeroing Reset

For most conventional CMOS APSs, the reset process is simply performed by a PMOS or NMOS transistor, and the fixed pattern noise is reduced by correlated double sampling(CDS) the output analog voltage. However, for TTFS_classic imagers, a digital signal is directly read out from a pixel, so it is impossible to apply CDS. By noticing that the major FPN in TTFS_classic imagers comes from the comparator random offset, we adopt a typical DC offset reduction method, autozeroing (AZ) [40], to reset the photodiode.

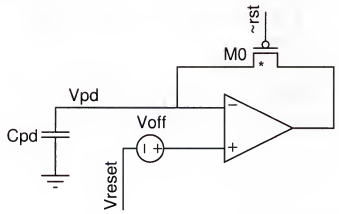


Figure 5–6: Autozeroing offset cancellation principle

During the reset phase, the comparator is disconnected from the signal path and connected as a unity-gain configuration as shown in figure 5–6. Assuming that the steady state has been achieved, the voltage V_{pd} obtained across the capacitor C_{pd} is

$$V_{pd} = V_{reset} + \frac{A}{1+A} V_{off} \tag{5.2}$$

After the reset phase, $M0$ is open and the comparator is connected again to the signal path. Thus, the offset V_{off} is stored on the capacitor if the comparator gain A is large enough. However, there is still an offset residue due to the finite opamp gain A, which is

$$V_{error,1} = -\frac{V_{off}}{1+A} \tag{5.3}$$

In addition to the finite gain, the charge injection from switch $M0$ and the clock feedthrough also cause errors. The total injected inversion charge is estimated to be

$$Q_{inj} = \alpha C_{ox} W_0 L_0 (V_{reset} + V_{th,p0}) \tag{5.4}$$

where α is the proportion of $M0$'s channel charge transferred to the capacitor C_{pd}, W_0/L_0 are the width and length of $M0$ respectively, and $V_{th,p0}$ is the threshold voltage of PMOS transistor $M0$ and has a negative value. Then the resulting error due to the charge injection is $\Delta V_{error,2} = Q_{inj}/C_{pd}$.

The clock feedthrough generates when $M0$ couples the clock transition to the sampling capacitor through its gate-source overlap capacitance C_{ov0}, and the resulting error is

$$\Delta V_{error,3} = V_{dd} \frac{C_{ov0}}{C_{pd} + C_{ov0}} \tag{5.5}$$

Then, the total residual offset equals

$$V_{os,res} = -\frac{V_{off}}{1+A} + \frac{\alpha C_{ox} W_0 L_0 (V_{reset} + V_{th,p0})}{C_{pd}} + V_{dd} \frac{C_{ov0}}{C_{pd} + C_{ov0}} \tag{5.6}$$

In general, a large opamp gain A is good for the offset residue reduction but degrades the comparator speed. With consideration of the worst threshold variations, the

opamp open loop gain is required to be at least 100 (see Appendix A). As the opamp unit bandwidth f_0 decides the autozeroing process settling time, a settling time of $t_{set} = 5\mu s$ demands $f_0 > 1/(2\pi t_{set}/7) \approx 224kHz$ to achieve 0.1% reset accuracy.

Reset speed has no specific requirement on the $M0$ design. Charge injection error and off-state leakage current, however, have great impacts on it. When $M0$ is biased in the weak inversion region, the drain current is given by [39]

$$I_{ds} = \mu_{eff}\, C_{ox} \frac{W_0}{L_0}(m-1)(U_T)^2 \exp\left(\frac{-V_{gs} + V_{th,p0}}{mU_T}\right)\left(1 - \exp(\frac{-V_{ds}}{U_T})\right) \qquad (5.7)$$

where μ_{eff} is the mobility of holes, $U_T = kT/q$ is the thermal voltage, and m is the body effect coefficient. It shows that subthreshold current exponentially depends on the threshold voltage and the gate-source voltage. So to achieve a low leakage current, a thick oxide transistor with a high threshold voltage is needed. In addition, the positive gate-source voltage during the comparison phase is also helpful to decrease the leakage current. The simulated leakage current is about 0.083fA at $T = 27°C$. It should be noted that the real leakage current may be larger than this number in that high working chip temperature will counteract this reduction as indicated in the exponential part of equation 5.7. A high threshold voltage is also useful to decrease the channel charge for the same V_{gs}, therefore reducing the charge injection.

In addition to the signal, noise is also stored on the capacitor as well. Autozeroing is equivalent to a high-pass filtering process, thus the low frequency noise, i.e., 1/f noise, is strongly reduced but at the cost of an increased noise floor due to aliasing the broadband white noise into the base band [40].

5.1.4 Comparator Design

A typical comparator utilized in A/D converters consists of a preamplifier and a latch and has two modes of operation: tracking and latching [41]. It, however, must work synchronously, which is not applicable in our design. In a TTFS_classic imager, there is no single global capture time, so we cannot decide when to latch the

comparison result. Since this comparator also works as an opamp during the reset phase, we need to consider the design requirements for both cases. The followings are the design criteria for this opamp/comparator:

1. Gain

   As explained in Appendix A, to ensure the success of autozeroing, the opamp gain should be greater than 100. During the comparison phase, the comparator essentially implements a one-bit comparison, so a high DC gain is not required in this phase. In summary, the opamp/comparator DC gain is required to be at least 100.

2. Open-loop Bandwidth

   In Chapter 4, we have shown that the comparison delay depends on the slew rate of the input signal, and the maximum delay is determined by the comparator open-loop bandwidth. In order to achieve good linearity in the reconstruction, we expect the comparison delay to be no larger than $2\mu s$, which corresponds to $f_{-3dB} = 80kHz$.

3. Close-loop Bandwidth

   As discussed earlier, the reset phase length requires the close-loop bandwidth is no less than 224kHz, which can be easily met.

4. Power and Size

   The comparator needs to be biased in the subthreshold region to minimize power consumption. Pixel size is one of our major concerns, since small pixel area means good spatial resolution. Thus the number of transistors must be minimized.

In our previous design, a cascaded opamp was used, which has a very high gain at the expense of speed. To boost speed, bias current must be increased, which is obviously contrary to the low power consumption requirement. So in our current design, two types of opamp, shown in figure 5–7 and figure 5–9 respectively, are under

our consideration mainly due to the speed and pixel size concerns. We will compare these two topologies in the following.

- **6-transistor Clamped Opamp**

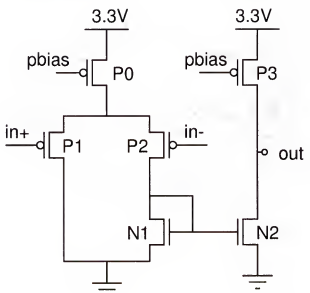


Figure 5–7: 6-transistor clamped opamp/comparator

1. Gain

   As required, all transistors are biased in the weak inversion region (or subthreshold region). The PMOS drain current is given in equation 5.7. When V_{ds} is much larger than a few U_Ts, the current can be generally simplified to

   $$I_{ds} = I_0 \exp\left((-V_{gs} + V_{th})/mU_T\right) \tag{5.8}$$

   where

   $$I_0 = \mu_{eff}\, C_{ox} \frac{W_0}{L_0}(m-1)(U_T)^2 \tag{5.9}$$

   It yields the transconductance

   $$g_m = \frac{\partial I_{ds}}{\partial(-V_{gs})} = \frac{I_{ds}}{mU_T} \tag{5.10}$$

Given $I_{ds,P1} = I_{ds,P2} = I_{ds,P3} = 0.5I_{ds,P0}$, all the transistors in figure 5–7 except $P0$ have the same transconductance g_m. Then the opamp gain is

$$\begin{aligned} A &= \frac{1}{2}\frac{g_m}{g_{ds,N2} + g_{ds,P3}} \\ &= \frac{1}{2}\frac{\frac{1}{mU_T}}{\frac{1}{V_{A,N2}} + \frac{1}{V_{A,P3}}} \end{aligned} \tag{5.11}$$

where $V_{A,N2}$, $V_{A,P3}$ are the early voltages of $N2$ and $P3$ respectively. The factor 0.5 comes from the fact that only half of the small signal current flows through transistor $P2$ to the output node. From CADENCE simulation with $I_{ds,P0} = 268nA$ and $I_{ds,P3} = 144nA$, a 40dB DC gain is achieved.

2. Speed

   Generally, the first pole is determined by the dominant time constant at the output node, which is

$$\tau = \frac{C_{out}}{g_{ds,N2} + g_{ds,P3}} \tag{5.12}$$

   where C_{out} is the total capacitance at the output node including the load capacitance and the parasitic capacitance of $N2$ and $P3$. Then the -3dB frequency is approximately

$$f_{-3dB} = \frac{1}{2\pi\tau} = \frac{1}{2\pi\frac{1}{g_{ds,N2} + g_{ds,P3}}C_{out}} \tag{5.13}$$

   With the same bias condition, the simulated f_{-3dB} is 170kHz.

3. Noise

   Since this clamped opamp has an asymmetric topology, the well known opamp noise model is not applicable. To see the noise performance, we need to analyze each individual component shown in figure 5–8. In the following derivations, we have assumed matching between transistor pair

$P1$ and $P2$ as well as in pair $N1$ and $N2$. It is very straightforward to analyze the noise contributions from $N1$, $N2$ and $P3$. For instance, the noise contribution from $N1$ to the output node is

$$V_{no} = (g_{m,N1} R_O) V_{n,N1} \tag{5.14}$$

where V_{no} is the output noise, $R_O = \dfrac{1}{g_{ds,N2} + g_{ds,P3}}$, and $V_{n,N1}$ is the equivalent noise source of transistor $N1$.

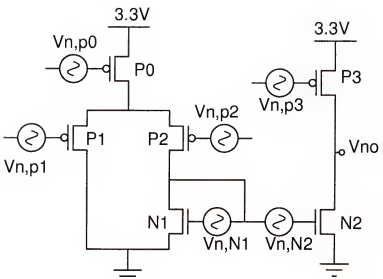


Figure 5–8: Diagram of 6-transistor opamp noise sources

For $P1$ and $P2$, notice that only half of the equivalent noise current of $P1$ is collected by the output node because of the asymmetric topology. Thus, we have

$$V_{no} = (\frac{1}{2} g_{m,P1} R_O) V_{n,P1} \tag{5.15}$$

Finally, the noise contribution from $P0$ to the output cannot be ignored as in symmetric opamps. Its noise current is divided into two components: one flows through $P1$ to the ground, and the other flows through $P2$ to

the output node. As a result, the noise contribution from $P0$ is given by

$$V_{no} = (\frac{1}{2} g_{m,P0} R_O) V_{n,P0} \tag{5.16}$$

Hence, the output total noise power will be

$$\begin{aligned} \overline{V_{no}^2} &= \frac{1}{4}(g_{m,P0}R_O)^2 \overline{V_{n,P0}^2} + \frac{1}{2}(g_{m,P1}R_O)^2 \overline{V_{n,P1}^2} \\ &+ 2(g_{m,N1}R_O)^2 \overline{V_{n,N1}^2} + (g_{m,P3}R_O)^2 \overline{V_{n,P3}^2} \end{aligned} \tag{5.17}$$

This output noise value can be referred back to get an equivalent input noise $\overline{V_{neq}^2}$ by dividing it by the gain $\frac{1}{4}(g_{m,P1}R_O)^2$, which results in

$$\begin{aligned} \overline{V_{neq}^2} &= 2\overline{V_{n,P1}^2} + \left(\frac{g_{m,P0}}{g_{m,P1}}\right)^2 \overline{V_{n,P0}^2} + 8\left(\frac{g_{m,N1}}{g_{m,P1}}\right)^2 \overline{V_{n,N1}^2} \\ &+ 4\left(\frac{g_{m,P3}}{g_{m,P1}}\right)^2 \overline{V_{n,P3}^2} \end{aligned} \tag{5.18}$$

If we size our devices such that $g_{m,P3}$, $g_{m,N1}$, $g_{m,P0} \ll g_{m,P1}$, we can minimize the noise contributions from devices $N1$, $N2$, $P3$ and $P0$. This can be accomplished by pushing those transistors into strong inversion by making $(W/L)_{P3,N1,P0} \ll (W/L)_{P1}$. In practice, we cannot increase their channel lengths arbitrarily due to the limited pixel size. Thus, these noise sources still dominate. Unlike that in strong inversion, the equivalent thermal noise source in weak inversion is approximately $\overline{V_n^2} = \frac{2kT}{g_m}$ [42], which yields

$$\overline{V_{neq}^2} = \frac{4kT}{g_{m,P1}} \left(1 + \frac{g_{m,P0}}{2g_{m,P1}} + 4\frac{g_{m,N1}}{g_{m,P1}} + 2\frac{g_{m,P3}}{g_{m,P1}}\right) \tag{5.19}$$

- **5-transistor Opamp**

  Figure 5–9 shows the other comparator under our consideration, which consists of a simple 5-transistor opamp shown inside the dashed box and an inverter

working as a second stage. The pseudo positive feedback is used to make the output node immune to the power supply noise.

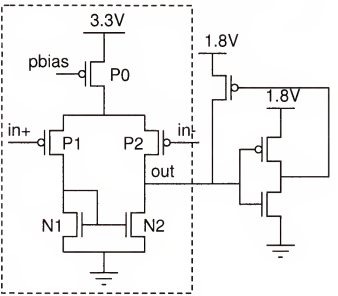


Figure 5–9: 5-transistor opamp/comparator

1. Gain and Speed

   Similar to the previous discussion, the 5-transistor opamp gain is

$$\begin{aligned} A &= \frac{g_{m,P1}}{g_{ds,N2} + g_{ds,P2}} \\ &= \frac{\frac{1}{m\hat{U}_T}}{\frac{1}{V_{A,N2}} + \frac{1}{V_{A,P2}}} \end{aligned} \tag{5.20}$$

   Since this 5-transistor opamp is a symmetric topology, there is no factor of 0.5 in the gain expression. With a bias current of 268nA, the simulated gain is 44dB, and the f_{-3dB} is 222kHz.

2. Noise

   The noise analysis for the 5-transistor opamp is quite simple. Assuming all the transistors work in the weak inversion region, the input referred

noise is

$$V_{neq}^2 = \frac{4kT}{g_{m,P1}} \left(1 + \frac{g_{m,N1}}{g_{m,P1}}\right) \tag{5.21}$$

Note that the noise contribution from $P0$ is ignored because of the symmetric topology. Obviously, this input referred noise is much smaller than that of 6-transistor opamp.

3. Residue

One potential problem with the 5-transistor opamp is the capacitor coupling through the overlap capacitor C_{ov_out} (see figure 5–10). After reset, the opamp output voltage V_{out} suddenly drops down from V_{reset} to almost 0V. This large swing of ΔV_{out} will introduce some error voltage expressed as

$$V_{error,A} = \frac{\Delta V_{out} C_{ov\_out}}{C_{pd} + C_{M1} + C_{ov\_out}} \tag{5.22}$$

The formula shows that a large C_{M1} can help reducing this error voltage. C_{M1}, however, cannot be made too large, since a very large C_{M1} will decrease the fill factor and also degrade the conversion gain.

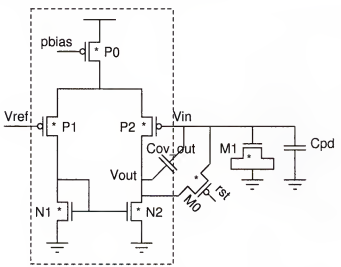


Figure 5–10: Diagram of the additional capacitor coupling in 5-transistor opamp

Comparing their performances, we prefer to use the 5-transistor opamp in that the high input referred noise floor makes the 6-transistor circuit unfavorable. The potential capacitor coupling problem with the 5-transistor opamp is not severe, since it can be reduced by increasing the MOS capacitor C_{M1}. By far, the overall offset residue after reset in terms of charge is given by

$$\begin{aligned} Q_{off,res} \approx & \left(-\frac{V_{off}}{A+1} + \frac{V_{dd}C_{ov0}}{C_{pd}+C_{M1}+C_{ov0}} \right. \\ & + \frac{\alpha C_{ox}W_0L_0(V_{reset}+V_{th,p0})}{C_{pd}+C_{M1}} \\ & \left. - \frac{V_{reset}C_{ov\_out}}{C_{pd}+C_{M1}+C_{ov\_out}} \right)(C_{pd}+C_{M1}) \end{aligned} \tag{5.23}$$

In our design, we used PMOS transistors as the input pair to save the pixel layout area. Theoretically, both PMOS and NMOS input pairs are able to provide the same signal swing if applying the autozeroing reset technique. However, since the lower bound of V_{ref} is limited to about 1V to achieve a constant capacitance for the MOS capacitor, a NMOS pair can be adopted to further increase the signal swing from 1.4V to 2V while trading off a large pixel layout area.

5.1.5 Digital Control Block Design

The digital control block inside each pixel consists of some simplified digital logic gates. No special design is required except one important issue, how to design the pixel interface to a column or row. To simplify the design of the large number of input OR gates per each row and column, a pseudo-CMOS logic is used instead. To obtain a fast switching speed, we need to optimally size the transistors $M6, M7$ and $M9$ in figure 5–1. Since $M6$ always switches on ahead of $M7$, the pull down delay is almost the same as that of a single NMOS. Typically, the always-on PMOS transistor has only limited driving current compared to NMOS transistors. The pull-down τ_{down} can be expressed as [39]

$$\tau_{down} = \frac{CV_{dd}}{2W_nI_{nsat}} \tag{5.24}$$

where I_{nsat} is the saturation current per unit width of NMOS $M7$, W_n is the width of $M7$, and C is the total output capacitance. Similarly, the PMOS pull-up delay is

$$\tau_{up} = \frac{CV_{dd}}{2W_p I_{psat}} \tag{5.25}$$

where I_{psat} is the saturation current per unit width of PMOS $M8$, and W_p is the width of $M8$. In this design, C mainly comes from the parasitic capacitance of NMOS $M6$, $M7$ and $M9$. Then,

$$C = kW_n C_0 \tag{5.26}$$

where k is the number of pixels in a row or column, and C_0 is the capacitance per unit width. Thus, we can simplify the switch delays as

$$\tau_{down} = \frac{kC_0 V_{dd}}{2I_{nsat}} \tag{5.27}$$

$$\tau_{up} = \frac{kW_n C_0 V_{dd}}{2W_p I_{psat}} \tag{5.28}$$

Note that W_n has no effect on the *pull-down* delay provided that W_n is large enough to ensure the correct logic. On the contrary, a large W_n shortens the *pull-up* delay due to an increased load capacitance. CADENCE simulation results in figure 5–11 have verified these observations. In addition, a small transistor width also helps to save layout area.

5.2 One Pixel Test

A test chip with one pixel was fabricated in TSMC $0.18\mu m$ digital CMOS technology. The chip has demonstrated the predicted functionalities. In the following, we will show some experimental results.

5.2.1 Light-Intensity-to-Time Transform

In a TTFS imager, the firing time is inversely proportional to the incident light intensity. To measure this light-intensity-to-time transform, we use three different light sources. With different neutral density filters covering the testing chip, a wide

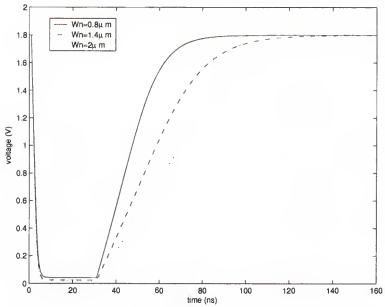


Figure 5–11: Pull-down and pull-up delay for different W_ns

range of luminance is provided. The light intensities in the testing environment are obtained by using a digital light meter LX-102. The measured transform is shown in figure 5–12, which demonstrates good linearity in the log domain as we expected. Limited by the available optics, the time response of very weak light intensities is not included in this measurement. We believe the firing time will saturate to 4.5s when the dark current becomes dominant. By putting the light source very close to the chip, we obtain the shortest firing time about $16\mu s$, which gives a measured dynamic range of 109dB. Since it is impossible for us to measure the actual light intensity under this circumstance, we did not include this data in figure 5–12 that only demonstrates about 90dB dynamic range.

5.2.2 Signal-Swing-to-Time Response

As we know, the firing time is not only a function of the incident light intensity but also a function of the signal swing. A reduced signal swing leads to a decreased firing time for the same light intensity. The measured signal-swing-to-time response in figure 5–13 shows good linearity in the large signal swing region. In contrast, the reset residue degrades the response linearity in the small signal swing region. After

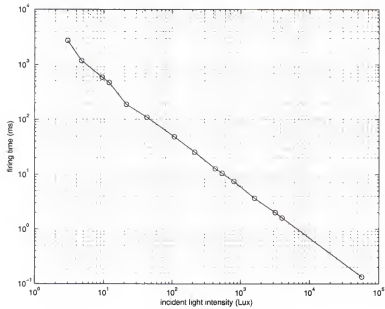


Figure 5–12: Light-intensity-to-time transform

calibrating, the linearity has been restored. The measured offset is around 50mV, which can be further reduced by increasing the capacitor C_{M1}.

5.2.3 Noise

The conventional CMOS imager noise measurement is performed in the analog domain, which is not suitable for time-based imager sensors. Therefore, we need to express the noise in terms of firing time variance. Suppose we have a measured signal i, which is expressed as

$$i = I + \Delta I \tag{5.29}$$

where I is the original clean signal or expected value of i, and ΔI is the noise component. Then, the actual firing time for the measured signal i is

$$t = \frac{Q}{i} = \frac{Q}{I}\left(\frac{1}{1 + \Delta I/I}\right) \tag{5.30}$$

For a signal with an acceptable SNR, we have $\dfrac{\Delta I}{I} \ll 1$ and

$$\frac{1}{1 + \Delta I/I} \approx 1 - \frac{\Delta I}{I} \tag{5.31}$$

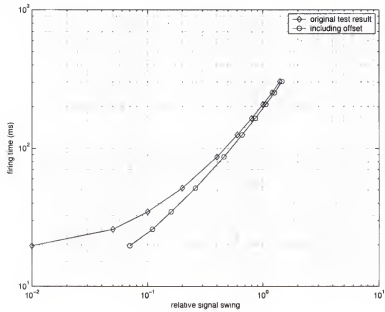


Figure 5–13: Signal-swing-to-time response

Following that, it yields

$$t = \bar{t}\left(1 - \frac{\Delta I}{I}\right) \tag{5.32}$$

and

$$\frac{\sigma_t^2}{(\bar{t})^2} = \frac{\overline{\Delta I^2}}{I^2} = \frac{\sigma_i^2}{(\bar{i})^2} \tag{5.33}$$

where $\bar{t} = \frac{Q}{I}$ and σ_t^2 are the mean and variance of the firing times over a sequence of frames respectively. So we can measure firing times to estimate analog noise. The major limitation to the accuracy of the measurement is the stability of the light source. Limited by current available optical instruments, we use a fluorescent light as the light source. We first set V_{reset} = 2V and V_{ref} = 1V. The measured firing times of 300 frames are shown in figure 5–14. The mean firing time is measured to be $11.04ms$, and the standard deviation is $52.7\mu s$, which gives a 46.4dB SNR. Another measurement is accomplished by adjusting V_{ref} to 1.4V. Since the well capacity is decreased, we expect the SNR will be somewhat degraded as we discussed in Chapter 3. We used the same light source and measured 300 frames again. This time, the mean firing time changes to $7.093ms$, which is almost 0.6 of the previous measurement

as expected. The standard deviation is 38.7μs, which corresponds to a degraded SNR of 45.2dB.

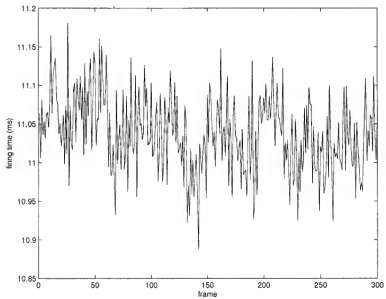


Figure 5–14: Temporal noise when $V_{ref} = 1$V

5.2.4 Conversion Gain

Conversion gain characterizes the signal generated per photoelectron and indicates the sensitivity of the sensor. An accurate determination of the conversion gain is also helpful to determine a photodiode's quantum efficiency [43]. The conversion gain is defined by

$$g = \frac{v}{n} = \frac{q}{C_{pd}} \tag{5.34}$$

where v is the signal voltage at the photodiode and n is the number of photoelectrons. B. Beechen and E. Fossum have proposed an excellent statistical method to determine the conversion gain based on an analog signal measurement [43]. For time-based image sensors, the conversion gain can be determined using the measured firing times by assuming the shot noise is the dominant noise source and obeys the Poisson statistics

[4]. Once the firing times are measured, the conversion gain can be determined by

$$\begin{aligned} g &= \frac{q}{C_{pd}} = \frac{q}{\frac{qN}{V_{reset} - V_{ref}}} \\ &= \frac{V_{reset} - V_{ref}}{N} \end{aligned} \tag{5.35}$$

where N is the integrated electrons, which can be estimated by $N = \left(\frac{\bar{i}}{\sigma_t}\right)^2$.

We captured the firing times using an Agilent 1693A Logic Analyzer. The sampling period is set to be $5ns$, which ensures the quantization noise is much lower than other noise sources. We choose $V_{ref} = 1\text{V}$ to guarantee the shot noise is the dominant source, since the photocurrent shot noise is proportional to the signal swing in time-based image sensors. From our previously measured results, we estimate the conversion gain to be

$$g = \frac{1}{\left(\frac{11.04}{0.0527}\right)^2} = 23\ \mu V/electron \tag{5.36}$$

It also gives

$$C_{pd} = \frac{q}{g} = 6.96fF \tag{5.37}$$

In general, a small well capacity or capacitance can be traded off for a large conversion gain.

5.3 Asynchronous Readout Design

A TTFS_classic imager is essentially a mixed-signal system with an analog circuitry inside each pixel and a digital readout circuitry surrounding the pixel array and dealing with the asynchronous readout scheme. Since the digital circuitry operates asynchronously, the standard Verilog or VHDL languages are not applicable to design it. Instead, our asynchronous readout circuitry is simulated by CADENCE SpectreS while considering the parasitics. The digital asynchronous readout circuitry mainly consists of a row arbiter tree, a column arbiter tree, a column latch, a column latch

control, a row interface and a throughput control block. We will detail each block in the following sections.

5.3.1 Arbiter

The arbiter cell shown in figure 5–15 is the central part of an asynchronous readout circuit and dealing with request collisions. An arbiter tree is built from two-input arbiter cells using the binary tree architecture. The arbiter used in the TTFS imager is not significantly different from that of Boahen [26].

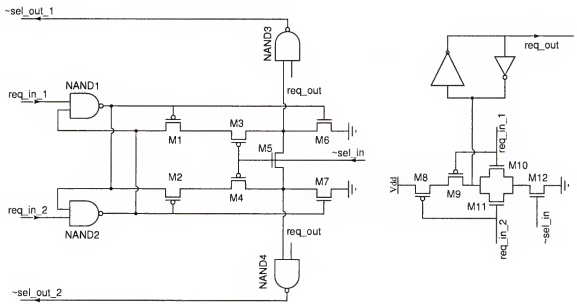


Figure 5–15: Schematic of arbiter cell

The truth table of the arbiter cell is given in table 5–1. It should be pointed out here that the logic combination of $req\_out_{init} = 0$ and $\sim sel\_in = 0$ does not exist, because without any request to the next stage, i.e., $req\_out_{init} = 0$, it is impossible to generate a response signal $\sim sel\_in = 0$ from the next stage. Whenever the current arbiter request to the next stage is approved, i.e., $req\_out_{init} = 1$ and $\sim sel\_in = 0$, an output signal logic is generated depending on the strength of the corresponding input signal. If a collision occurs, two inputs will compete to ensure only one output, either $\sim sel\_out\_1$ or $\sim sel\_out\_2$, is activated.

Table 5–1: Truth table of arbiter cell

Initial Condition	Input			Output		
req_out_{init}	$\sim sel\_in$	req_in_1	req_in_2	req_out	$\sim sel\_out\_1$	$\sim sel\_out\_2$
0	1	0	0	0	1	1
0	1	0	1	1	1	1
0	1	1	0	1	1	1
0	1	1	1	1	1	1
1	0	0	0	0	1	1
1	0	0	1	1	1	0
1	0	1	0	1	0	1
1	0	1	1	1	0 1	1 0
1	1	0	0	0	1	1
1	1	0	1	1	1	1
1	1	1	0	1	1	1
1	1	1	1	1	1	1

5.3.2 Latch Cell and Latch Control

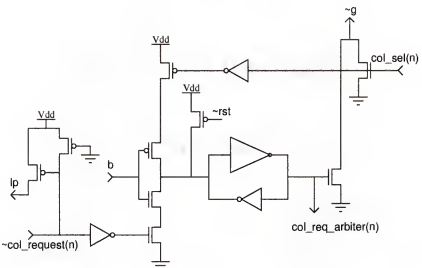
To boost the throughput of the asynchronous readout, latch and latch control circuitries are incorporated. Their schematics are shown in figure 5–16 [26]. Note that there is one latch cell for each column and one single latch control circuit for the whole array. The latch cells and the latch control block need to cooperate together properly to generate a correct logic. They operate as follows:

1. The initial condition is obtained in the reset period by activating the control signal $\sim rst$ and deactivating $\sim col\_request$(n), i.e., $\sim rst$=0 and $\sim col\_request$(n)=1. Then we have $col\_req\_arbiter$(n)=0, col_sel(n)=0 (since $col\_req\_arbiter$(n)=0, no response comes from the column arbiter), $l_P = 0$, $\sim g = 1$, $\sim row\_sel\_enable$=0, $b = 1$, and $latch\_data\_ready$=0.

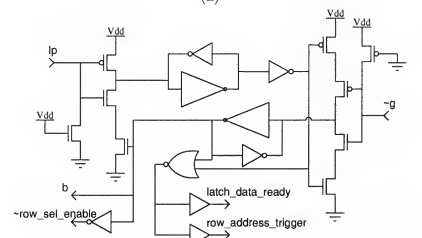
2. Once a column request comes in, i.e., $\sim col\_request(n) = 0$, it pulls up l_P and sends out a request signal to the column arbiter by setting $col\_req\_arbiter(n) = 1$, which in turn pulls down $\sim g$. Then $\sim g = 0$ and $l_P = 1$ lead to $b = 0$ and latch all the column requests by cutting off the input path. Simultaneously, $\sim row\_sel\_enable = 1$ disables the pixel column request function and guarantees no new column request is generated. At the same time, the active signal $latch\_data\_ready$ and $row\_select(m)$ produce a control signal $disable(m)$, which

together with $col\_req(n)$ disables the latched pixels in the selected row (see figure 5–1). In addition, $row\_address\_trigger = 1$ makes the row address encoder update the output row address.

3. When all the column request signals are processed and withdrew after receiving the feedback signal $col\_sel(n) = 0$, $\sim g$ returns to high. Since no new $\sim col\_request(n)$ arrives, $l_P = 0$ brings all the control signals back to their initial conditions, and all the latch cells are open again for the new upcoming $\sim col_request(n)$s.



(a)



(b)

Figure 5–16: Schematics of latch cell and latch control. (a) latch cell. (b) latch control.

5.3.3 Row Interface

The row interface shown in figure 5–17 has two functions, i.e., block unwanted $row\_request(m)$ signals and decide whether a row can be selected by the row arbiter. $\sim isolation$ is the same signal that we used in the pixel design. During the reset period, the uplink path to the row arbiter is disabled, and $row\_req\_arbiter$ is initially set to 0. Similarly, the downlink path from the row arbiter is disabled by setting $\sim row\_sel\_enable = 1$ once there is an occupied latch cell.

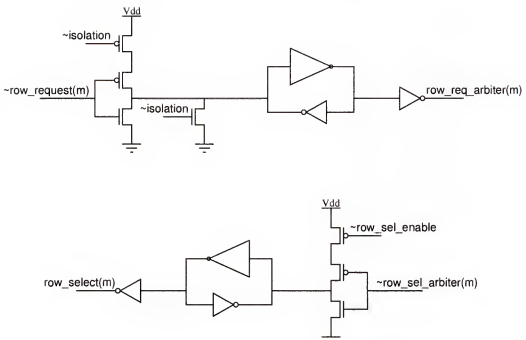


Figure 5–17: Schematic of row interface

5.3.4 Throughput Control

Usually, a communication channel between neuromorphic chips using address-event representation needs a *request* line and an *acknowledge* line. A typical arbiterred communication channel adopts a *handshake* protocol, i.e., a new transmission cannot start until the *acknowledge* signal for the last transmission is received. A *handshake* method inevitably increases the time-overhead of one arbitration by adding an *acknowledge* time period into the communication cycle. Though many methods

have been proposed to shorten one complete cycle [44], it is still too long for our TTFS_classic imager if considering the reconstruction errors, which will be discussed in the next chapter. Since randomly lost pixels can be reconstructed by a spatial interpolation method, we can relax this tight requirement on communication channel design by simply deleting the *acknowledge* line. Instead, we designed a readout control block shown in figure 5–18, which is controlled by an external clock to make testing easier. Two non-overlapping clocks $phase1$ and $phase2$ control the transmission gates, so at most one transmission gate is open at any time. The signal $col\_sel\_arbiter$(n) comes from the column arbiter, the signal col_sel(n) goes to the latch cell, and the signal $col\_encoder\_in$(n) is the input to the column address encoder.

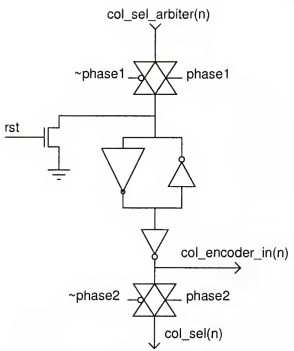


Figure 5–18: Schematic of throughput control

5.4 Layout

A prototype TTFS_classic imager was implemented in TSMC $0.18\mu m$ digital technology and packaged in a standard 84-pin PGA available through MOSIS. The total die size is $5mm \times 5mm$. The layout is shown in figure 5–19, which has a 128x128

pixel array sitting in the center of the whole chip surrounded with guard rings. To reduce the substrate coupling, the guard rings are shorted to a separate pair of power supply and ground. In addition, analog, digital and pad circuitries use different power supplies to avoid corrupting the sensitive analog signals, and each of them is formed by multiple input pins to reduce the equivalent parasitic inductance. Also, to prevent light-induced currents from affecting the analog circuitry or causing latchup, the sixth layer metal is used everywhere except over the photosensitive node.

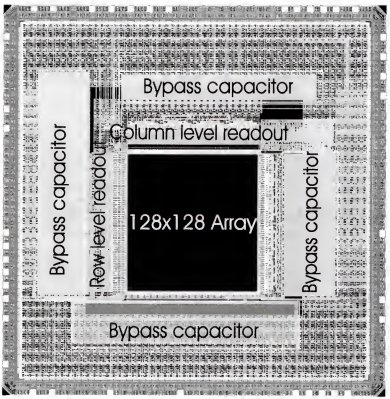


Figure 5–19: Layout of the prototype TTFS_classic

Numerous on-chip bypass capacitors are also included to reduce SSN. One potential problem with on-chip bypass capacitors is the resonance oscillation in the power distribution network [45]. In our design, we estimate the total power supply parasitic inductance L_P to be less than 2nH resulted from multiple power supply pins and mutual inductance. The maximum on-chip bypass capacitance C_{by} is simulated to be about 3nF. It gives us the minimum resonance frequency about

$w_r = \sqrt{\frac{1}{2L_P C_{by}}} \approx 289MHz$. The clock we used in the test is below 50Mhz, which is much less than the minimum resonance frequency. Thus, the unexpected resonance oscillation in the power supply network did not happen in either our CADENCE simulations or real-time image capture test. If a much faster clock is used in the test, an on-chip parasitic resistance needs to be included in series with the on-chip bypass capacitor [45]. In addition, we optimize the digital output buffer to limit its driving current to reduce the SSN amplitude while ensuring its capacity to drive a standard 15pf load within 2ns. To further reduce the SSN, off-chip resistors are added in series with each address output pin.

5.5 Testing and Characterization of 128x128 Sensor Array

A small printed circuit board (PCB) shown in figure 5–20 is built to integrate the TTFS_classic imager, voltage regulators, bias generators and connectors. A constant reference voltage is generated on the PCB board. Once the timing information is collected by an Agilent 1693A logic analyzer working in the asynchronous sampling mode, a simple MATLAB program reconstructs the captured images.

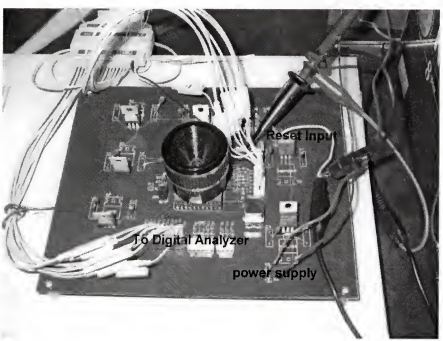


Figure 5–20: Experimental setup

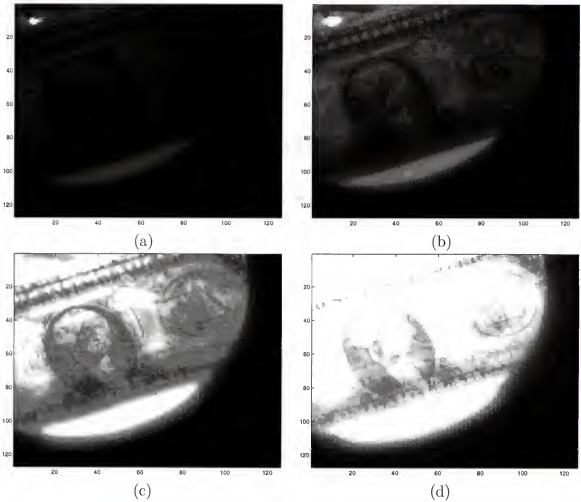
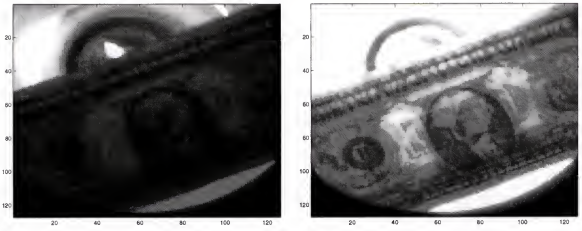


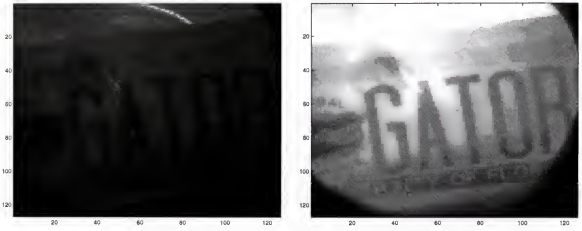
Figure 5–21: A 98dB dynamic range image. (a) scale=1. (b) scale=16. (c) scale=64. (d) scale=128.

The high dynamic range test is first setup with a one dollar bill taped over an incandescent lamp and impinged onto our imager. We captured several high dynamic range images by adjusting the height of the object. A 98dB dynamic range image is shown in figure 5–21. If we map the brightest pixel to the maximum discrete display level, only the center of incandescent bulb appears in left upper corner of figure 5–21(a), and the shape of the lamp appears in the lower region of the image. With the same captured image being scaled by different levels, more details of the dark region show up. In figure 5–21(b), the obscure image of the dollar bill is present, whereas the details of the dollar bill is clearly displayed in figure 5–21(c). Note that part



(a) (b)

Figure 5–22: A 75dB dynamic range image. (a) bright region details. (b) dark region details.



(a) (b)

Figure 5–23: Gator logo 1. (a) bright region details. (b) dark region details.

of the filament also shows up in the left upper corner of this figure. As illustrated in figure 5–21(d), if the scaling factor is too big, most of the pixels saturate and are barely visible except for the portrait of president George Washington. Another high dynamic range image of 75dB is displayed in figure 5–22, which contains more detailed information of both the lamp holder and the filament. These two images have a 20dB dynamic range difference, which is due to the fact that the darkest pixel in the first capture is much weaker than that in the last one. The experimental results are much beyond the dynamic range limitation of up-to-date commercial CMOS APSs.

Figure 5–24: Gator logo 2

A more complicated experimental setup is necessary in order to get much higher dynamic range images, which is not feasible considering our current available test equipments. Here, we also give two other captured images of two different gator logos in figure 5–23 and figure 5–24 respectively. Again, after scaling the image, more details in the dark region show up in figure 5–23(b), e.g., part of the word "UNIVERSITY OF FLORIDA" at the bottom and half of the gator head at the left. Note that all the displayed images here are collected by a single capture and without postprocessing. In each capture, though some missed firing pixels are present, they can be reconstructed from a spatial interpolation method.

The maximum power consumption of the whole system is measured to be less than $8.7mW$ when the sensor operates at $30frames/sec.$ For the fixed pattern noise measurement, a strict uniform light source is required. Unfortunately we do not have such optical instrument, and we used a typical fluorescent lamp from the ceiling as the uniform light source instead to measure the firing times without a lens. In this way, the FPN is estimated to be about 2.3%.

5.6 Summary

In this chapter, we have discussed the detailed circuit design of the TTFS_classic image sensor. The characterization is done at both pixel and array levels in terms of power consumption, light-intensity-to-time transform, signal-swing-to-time response, noise, conversion gain and dynamic range. The demonstrated test results are very promising. Due to the lack of the necessary optical equipment, the FPN is only estimated in a typical room light environment. The performance of the TTFS_classic imager is summarized in table 5–2. We did not observe any reconstruction error due to the readout collision from our prototype TTFS_classic imager captured images. However, with the TTFS_classic scaling up to a large size array, the reconstruction error might be prominent. We will deal with this issue in the next chapter.

Table 5–2: The performance of the TTFS_classic imager.

Technology	0.18 μm TSMC Digital CMOS
Supply Voltage	3.3V(analog), 1.8V(digital and pad)
Transistors per pixel	24
Array size	128×128
Pixel size	$12.4\mu m \times 12.1\mu m$
Photosensitive area	$2.99\mu m \times 2.97\mu m$
Die size	$5mm \times 5mm$
Power dissipation	8.7mW max @30 frames/second
Dark current	$\leq$1.04nA/cm^2 @room temperature
Conversion Gain	23 μV/electron
Dynamic Range(single pixel)	109dB(measured) >133dB(theory)
Dynamic Range(array)	98dB(measured) >113dB(theory)
SNR(measured)	46.4dB with signal swing=1V
FPN(estimated)	2.3%
Package	PGA84M

CHAPTER 6
MODIFIED TTFS IMAGERS

In this chapter, we will investigate the potential readout errors for a large array size TTFS imager. Several modified TTFS imagers are proposed here to reduce the readout errors.

6.1 Performance Limits for TTFS_classic

One unique issue with the asynchronous readout is the situation when many pixels may be under similar illumination. These pixels will send out *request* signals within a short period. If the readout circuit is not fast enough to output all the *request* signals in real time, some amount of delay will be inevitably introduced resulting in some temporal errors for the pixel illumination. For each illumination, the relative error can be expressed as

$$RE = \frac{I - I_r}{I} = \frac{t_r - T}{t_r} = \frac{\Delta t}{T + \Delta t} \tag{6.1}$$

where I and I_r are the original photocurrent and reconstructed photocurrent respectively, $\Delta t = t_r - T$ is defined as the time delay of the output pulses. The above equation indicates that for the same amount of time delay Δt, the error is more serious for a high illuminance pixel, where the firing time T is very small. Also, this problem is more serious for larger size images, where the probability that many pixels are firing in a short period is higher. According to our analysis based on TSMC $0.18\mu m$ digital CMOS process results, the acceptable shortest integration time is $10\mu s$ for QCIF (144×176) image format, and it is $100\mu s$ for a 480×720 size image.

The readout time delay is introduced at two places, i.e., the column arbitration for reading out request pixels in the same row, and the time needed for disabling pixels, row arbitration and latching column *requests*. Two obvious directions for readout

time delay reduction are either to improve the readout throughput by decreasing the circuit delay or to modify the system architecture to reduce the collision possibility.

To evaluate the severity of this issue, a MATLAB-based TTFS imager simulator was built. Critical timing information is extracted from SPICE simulations by considering the worst case. The four high dynamic range (HDR) images, namely *nave, groveC, rosette, and vinesunset*, used in the simulation are from Paul Debevec's graphics research group at the University of Southern California [46]. These four images use a floating-point representation and have dynamic ranges from 88dB to 168dB. The subsampled images are slightly larger than QCIF (144×176) size images, which are commonly used in hand-held devices. The statistics of these images are summarized as below.

Table 6–1: Four 480×720 and four 160×180 HDR images used in MATLAB simulation.

Images with a size of 480×720				
Image Name	nave	groveC	rosette	vinesunset
Minimum Value	1.69×10^{-5}	5.53×10^{-4}	2.63×10^{-5}	1.44×10^{-3}
Maximum Value	4.27×10^{3}	8.80×10^{2}	8.28×10^{1}	3.61×10^{1}
Dynamic Range(dB)	168	124	130	88
Images with a size of 160×180				
Image Name	nave	groveC	rosette	vinesunset
Minimum Value	1.69×10^{-5}	9.13×10^{-4}	2.63×10^{-5}	1.58×10^{-3}
Maximum Value	2.59×10^{3}	4.93×10^{2}	7.43×10^{1}	3.31×10^{1}
Dynamic Range(dB)	164	115	129	86

Table 6–2: The simulated mean relative errors introduced by the TTFS_classic imager for video mode applications

Image	*nave*	*groveC*	*rosette*	*vinesunset*
160×180	1.844×10^{-5}	1.887×10^{-5}	1.400×10^{-6}	1.121×10^{-2}
480×720	1.7772×10^{-3}	1.90×10^{-2}	1.8365×10^{-4}	5.9142×10^{-1}

For the 160×180 size images, the simulated mean relative errors (MREs) introduced by a TTFS_classic imager are listed in table 6–2. These MREs correspond to SNRs of more than 40dB if time-delay introduced errors are regarded as the only

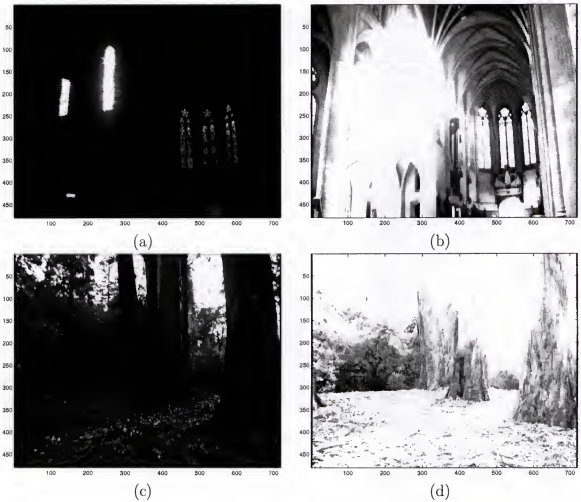


Figure 6–1: Images *nave* and *groveC*. (a) *nave* (bright part). (b) *nave* (dark part: $\times 100$ brighter for display). (c) *groveC* (bright part). (d) *groveC* (dark part: $\times 20$ brighter for display).

noise component. Considering typical CMOS APS imagers have a maximum SNR around 40dB, the above errors are not significant. On the contrary, for very large size images, the TTFS_classic imager cannot handle the large throughput, thus introducing more reconstruction errors (see table 6–2). For instance, image *vinesunset* has a large portion of pixels firing shortly after reset, e.g., within $100\mu s$, and its reconstruction error is inevitably higher.

The above numerical simulation shows the asynchronous readout delay limits a large size TTFS_classic imager dynamic range at the high end of illuminance (or

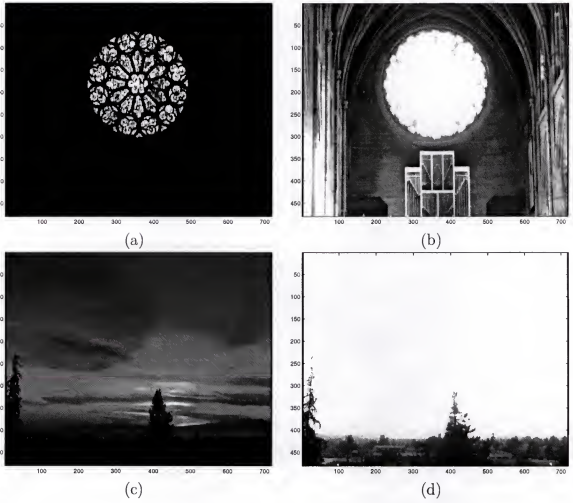


Figure 6–2: Images *rosette* and *vinesunset*. (a) *rosette* (bright part). (b) *rosette* (dark part: $\times 100$ brighter for display). (c) *vinesunset* (bright part). (d) *vinesunset* (dark part: $\times 10$ brighter for display).

equivalently at the low end of firing time). Thus, more efficient designs are necessary to reduce the readout errors.

6.2 TTFS with Rolling Shutter

In the previous section, the MATLAB simulation has demonstrated that the readout delay error of the TTFS_classic is not significant for medium size images with dynamic range of more than 100dB. However, the asynchronous readout still limits the achievable dynamic range by the shortest firing time (e.g., $100\mu s$) for large size images. If a large number of pixels fire shortly after global reset, the high collision rate will make the readout errors more prominent. To avoid a high collision rate for this case,

we propose a modification to the TTFS_classic architecture using a *rolling shutter* technique. The new architecture (TTFS_RS) has a separate reset signal for each row instead of a single global reset. For natural images, neighboring pixels intensities are strongly correlated [47], so the firing times of neighboring rows tend to occur close in time. De-correlating the firing times decreases the probability of simultaneous readout requests, therefore reducing the collision rate. The rolling shutter spreads out the firing times of neighboring rows. An alternative implementation would be to place filters with differing attenuating factors over the pixel array. This would also decorrelate neighboring pixel values but still use the same global reset signal. The attenuating values would have to be known or calibrated for each pixel in order to reconstruct the image.

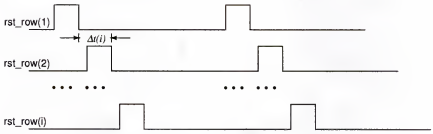


Figure 6–3: Rolling shutter pattern for still mode applications

For still mode applications with a constant reference voltage scheme, we can adopt a typical rolling shutter pattern shown in figure 6–3, and there is no specific modification on the scene reconstruction. For video mode applications, there are two alternatives to provide the reference voltage, one D/A conversion per row or one D/A conversion per chip. For the former case, the typical rolling shutter pattern is still feasible, while a special rolling shutter pattern must be generated for the latter case. If the whole array shares one D/A conversion, the rolling shutters must have the same rising edge or starting time (see figure 6–4). A simple digital circuit shown in figure 6–5 can be used to generate this special rolling shutter pattern. Every k rows, e.g., 96, are grouped together to share one rolling shutter pattern. Inevitably, this

technique will add some burden on reconstruction. For *row m*, if a global piece-wise linear reference voltage pattern is applied, the time stamp $T_0(m)$ when the reference voltage starts to change is

$$T_0(m) = T_0(0) - (m \bmod k) \times \Delta t_{shutter} \tag{6.2}$$

where $\Delta t_{shutter}$ is the unit shutter delay. And the corresponding frame period becomes

$$T_{frame}(m) = T_{frame}(0) - (m \bmod k) \times \Delta t_{shutter} \tag{6.3}$$

If the time stamp for the received pixel(m, n) is $t_r(m, n)$, the reconstructed photocurrent can be calculated by

$$I_r(m,n) = \begin{cases} \dfrac{Q_{well}}{t'_r(m,n)} & \text{if } t_r(m,n) < T_0(m) \\ \dfrac{Q_{well}}{t'_r(m,n)} \cdot \dfrac{T_{frame}(m) - t'_r(m,n)}{T_{frame}(m) - T_0(m)} & \text{if } t_r(m,n) > T_0(m) \end{cases}$$

where $t'_r(m,n) = t_r(m,n) - (m \bmod k) \times \Delta t_{shutter}$, and Q_{well} is the well capacity given by $Q_{well} = (V_{reset} - V_{ref})\, C_{pd}$.

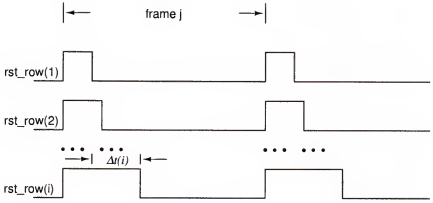


Figure 6–4: Rolling shutter pattern for video mode applications

No matter what rolling shutter pattern is used, captured images will demonstrate scattered firing times. Here we take image *vinesunset* as an example to show how efficient the rolling shutter can scatter the firing times. Compared to that for a global

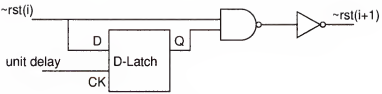


Figure 6–5: Rolling shutter generator for video mode applications

reset in figure 6–6, the firing time histogram for the rolling shutter reset demonstrates a large variation as expected.

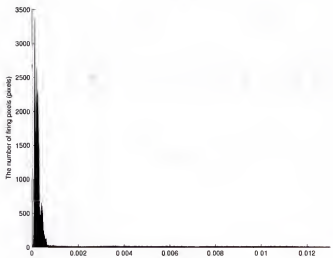


Figure 6–6: Simulated firing time histogram of *vinesunset* for the TTFS_classic imager

To gauge the quality of a TTFS_RS imager, we investigate MREs for those HDR images with 480×720 array size. The simulated MREs listed in table 6–3 demonstrate that the TTFS_RS imager successfully decreases the reconstruction error by spreading out the firing times. For still mode applications, a longer shutter delay is preferred. However, the rolling shutter delays between pixel resets must be kept short in dynamic scenes to prevent motion artifacts. To help visualization, we include the reconstructed *vinesunset* images in figure 6–8. Obviously, the reconstructed image by the TTFS_classic has more prominent errors in the bright part, whereas the TTFS_RS has successfully reduced the magnitude of the errors and no visible artifacts remain in the reconstructed image.

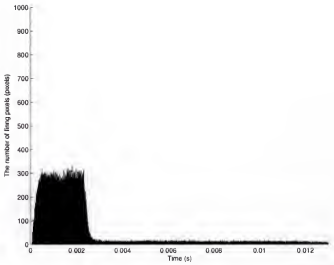


Figure 6–7: Simulated firing time histogram of *vinesunset* for the TTFS_RS imager.

Table 6–3: The simulated mean relative errors introduced by the TTFS_RS imager for video mode applications

Image	*nave*	*groveC*	*rosette*	*vinesunset*
$delay = 0\mu s$	1.7772×10^{-3}	1.90×10^{-2}	1.8365×10^{-4}	5.9142×10^{-1}
$delay = 20\mu s$	4.1360×10^{-6}	2.0892×10^{-3}	4.3259×10^{-5}	6.6035×10^{-3}
$delay = 35\mu s$	3.5200×10^{-6}	1.7239×10^{-3}	5.3687×10^{-5}	1.2677×10^{-3}
$delay = 50\mu s$	3.7302×10^{-6}	1.0456×10^{-3}	6.8610×10^{-5}	4.2494×10^{-4}

6.3 Hybrid TTFS Imager

Biological vision and electronic image acquisition share some common principles, and local memory is one of the common features [48]. A digital pixel sensor (DPS) with local memory has been designed by the Smart Image Sensor Group at Stanford University [49], which shows the feasibility of implementation of local memory in silicon. The potential high collision rate of asynchronous readout makes the TTFS_classic unfavorable for a very large size array. However, this problem can be avoided by incorporating local memory. To achieve a fine A/D conversion for a wide dynamic range of 120dB, at least 20 bits of local memory are needed for conventional time-based methods, which will dramatically increase the pixel layout area and degrade the sensor's spatial resolution for a typical CMOS process. The thin

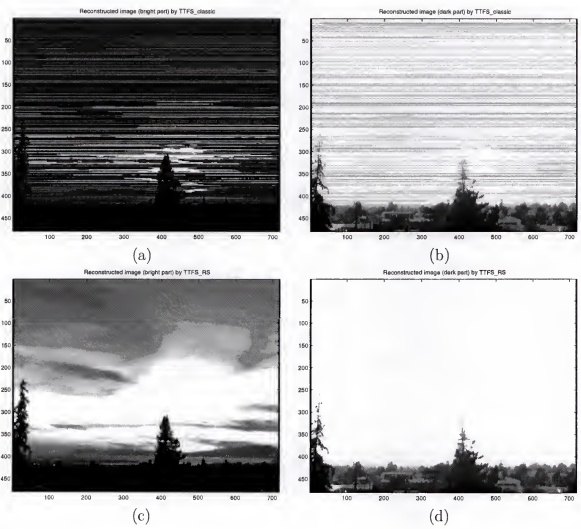


Figure 6–8: Image *vinesunset*. (a) bright part of the reconstructed image by the TTFS_classic. (b) dark part of the reconstructed image by the TTFS_classic. (c) bright part of the reconstructed image by the TTFS_RS. (d) dark part of the reconstructed image by the TTFS_RS.

film on ASIC (TFA) technology is probably the best means to implement this architecture, since it vertically integrates an amorphous silicon detector and a crystalline application-specific circuit for pixel readout and optional signal processing [50]. However, introducing special layers into a conventional process will tremendously increase the cost of fabrication. Therefore, for an image sensor fabricated in a conventional CMOS process, we have to decrease the local memory area to favor the spatial resolution. Luo [31] has proposed a two-degree-of-freedom quantization technique, which

can use 8 bits to represent over 90dB high dynamic range images. It is in fact a pseudo-log compression of wide dynamic range image with sacrificing the quantization noise. However, it greatly improves the possibility of using local memory for time-based image sensors. Here, we adopt this two-degree-of-freedom quantization principle and propose a novel TTFS imager with local memory, called TTFS_LM.

6.3.1 Principle of TTFS_LM

The pixel schematic of this novel sensor is shown in figure 6–9, which has a photodiode, a comparator, a level-down shifter and a 9-bit 3T-DRAM. Unlike Stanford's DPS, this novel image sensor works in a continuous mode to sense the input transition, thus a photodiode instead of a photogate is used. Note that the positive and negative signs in the opamp/comparator block denote the inputs only for the opamp, and the signs should be switched in the comparator schematic (see figure 6–11 for details). We apply the varying reference voltage to the positive input of the opamp and generate the digital value in an external counter. By carefully adjusting the reference voltage, we could implement uniform or nonuniform quantization. The whole circuit has two phases, time-to-first-spike (or saturation sensing) phase and conventional fine A/D conversion phase. The transistors labelled with $*$ are thick oxide 3.3V transistors, which are used to achieve a large input signal swing and a low leakage current.

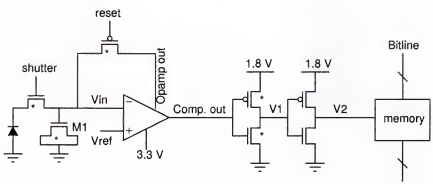


Figure 6–9: Pixel schematic of the TTFS_LM imager

During the time-to-first-spike period or phase 1, the photocurrent discharges the photodiode and V_{in} continuously drops down until it reaches a constant reference voltage $V_{ref} = V_{min}$. Then the comparator flips, $V1$ goes to high, and $V2$ goes to low. Simultaneously the digital value generated by the external counter is latched in the local memory. The most significant bit (MSB) indicates that the digital value latched here is for phase 1. The digital value can be generated in terms of uniform or nonuniform quantization. If using uniform quantization, the achievable dynamic range is only $6 \times 9 = 54dB$, while for nonuniform quantization (such as log compression), the dynamic range can be enhanced over 100dB.

During the fine A/D conversion period, i.e., phase 2, the remaining voltage of the photodiode will be firstly stored on $M1$ by closing the *shutter*. At this moment, we apply a ramp voltage to V_{ref}. The beginning ramp voltage is lower than the expected lowest voltage at the sense node, which causes $V2$ to raise and enables the memory to load the gray code values generated by the external counter. Next, the ramp voltage linearly increases until beyond the reset voltage. When the ramp exceeds the sense node voltage, $V2$ goes low, and the pixel memory latches the corresponding gray code. The 8-bit gray code is stored in the LSBs of the local memory, while MSB=0 represents the fine A/D conversion.

Figure 6–10 illustrates the principle of this novel TTFS_LM imager. Based on photocurrent amplitudes, the memory latches different digital outputs, a log-compression code or gray code for the single slope A/D conversion. Of course, we can replace the log-compress code with other codes, e.g., a uniform quantization code. It clearly shows that the local memory latches *only once* for each capture or frame. No readout delay or confusion generates, hence the TTFS_LM solves the potential collision problem with the TTFS_classic.

In summary, this novel design has the following advantages:

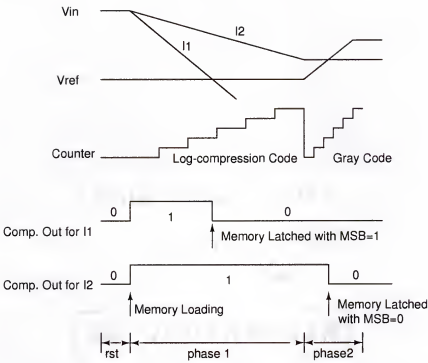


Figure 6–10: Principles of the TTFS_LM imager

1. Be able to decide the optimal shuttering time. The first 2^8 levels can be dedicated to obtain a rough statistics of pixel firing times and decide the optimal shuttering time.

2. Extend the dynamic range for each capture or frame. How much dynamic range extended depends on the quantization method adopted in the phase 1.

3. Be able to realize a multiple sampling technique, which can be implemented by assigning individual capture times or shuttering times.

4. Be able to capture a high dynamic range scene with fewer captures than a conventional multiple sampling technique.

5. Have a simpler V_{ref} pattern than that for Luo's method.

6. Enjoy better SNR, which will be discussed in section 6.3.4.

6.3.2 Pixel Design

Since both nonuniform and uniform quantization can be used in this system, the digital CDS technique [49] is not applicable here. Instead, we use an autozeroing technique to reset the photodiode and reduce the offset FPN simultaneously. Thus

the comparator is required to work also as an opamp. To achieve a small layout area, a relative high gain and a large bandwidth, we use a two stage comparator topology shown in figure 6–11, where the first stage is a conventional 5-transistor opamp, and the second stage is made up of $N3$ and $P3$. All these transistors are thick oxide transistors. The followings are the design criteria for this opamp/comparator:

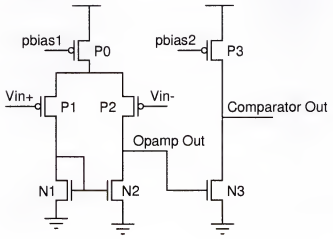


Figure 6–11: Comparator/opamp in the TTFS_LM

1. Assume the maximum ADC resolution to be m=8 bits and the trigger point for the following level-down shifter is 1V. If the input voltage is 2V, thus the comparator gain is required to be at least $1/2 \times 2^8 = 128$.

2. Suppose the total A/D conversion time is $500\mu s$. Then each comparison time is $500\mu s/256 = 1.95\mu s$, which means the comparator bandwidth must be at least 81.5 kHz.

3. If the maximum comparator introduced delay is required to be less than $2\mu s$, then the lower bound of opamp bandwidth is 79.5 kHz, which is less than 81.5 kHz.

4. Minimize power consumption.

Table 6–4 lists the performance of this topology from CADENCE simulations, which has satisfied all the criteria. The whole circuit operates in the weak inversion region to save power consumption and achieve a high comparator/opamp DC gain.

Table 6–4: Performance of the proposed opamp/comparator

Opamp /Comparator gain	44dB/75.8dB
Opamp/comparator f_{-3dB}	114kHz
Phase margin	79°
Total current	300nA

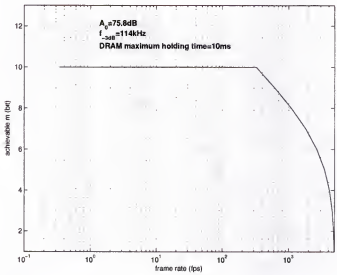


Figure 6–12: Achievable bit resolution vs. frame rate

The maximum achievable ADC resolution is limited by the opamp bandwidth. There are two ways to improve it. One way is to increase the bias current, thus improving the bandwidth. Of course, this will increase the power consumption, simultaneously decrease the opamp gain and affect the FPN reduction. The other way is to lengthen the A/D conversion period, which would lower the frame rate. Suppose the maximum DRAM holding time is 10ms and the fine A/D conversion spends half of the holding time. We plotted the simulated achievable ADC resolution in figure 6–12, which shows that a higher frame rate leads to a lower ADC resolution. As discussed, both the comparator gain and the maximum DRAM holding time limit the upper bound of achievable ADC resolution. The 3T-DRAM is shown in figure 6–13. A typical design has achieved a maximum data hold time of 10ms [49], which gives an ADC resolution of $m = \lfloor (10ms/2) \cdot 2\pi \cdot f_{-3dB} \rfloor = 10$. Overall, there are 38

transistors in each pixel. We expect each pixel will occupy a $12\mu m \times 12\mu m$ area and the fill factor is around 15% for a typical $0.18\mu m$ digital CMOS technology.

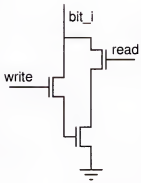


Figure 6-13: Schematic of 3T-DRAM in the TTFS_LM

6.3.3 Clock Pattern for Uniform Quantization in Phase 1

To achieve the uniform quantization of photocurrent, a special clock pattern is required. Suppose the shutter is closed at T_0, and in the following period, the single slope A/D conversion is performed. Thus, the maximum nonsaturated photocurrent in phase 1 is

$$I_0 = \frac{Q_{well}}{T_0} \tag{6.4}$$

where Q_{well} is the available well capacity. Assume a 9-bit pixel level DRAM is integrated, then the photocurrent resolution is

$$\Delta I = \frac{Q_{well}}{T_0 \cdot 2^8} \tag{6.5}$$

In order to achieve the same quantization resolution in the saturation sensing period with a fixed reference voltage, we have to adjust the timing for the external ramp generation. For $I_k, k = 1, \cdots, 256$, the individual timing would be

$$\begin{aligned} t_k &= \frac{Q_{well}}{I_0 + k\Delta I} \\ &= \frac{T_0}{1 + k/256} \end{aligned} \tag{6.6}$$

This expression not only gives us the specific timing information but also provides us the speed requirement for an external clock to achieve enough A/D conversion accuracy.

6.3.4 Novel Multiple Sampling Technique

A conventional multiple sampling scheme is to sample the pixel output at exponentially increasing exposure times, T, $2T$, $\cdots$, 2^kT. Since the TTFS_LM pixel is able to enhance each capture dynamic range by a factor of 2 if uniform quantization is used in phase 1, the exposure time can be set to be exponentially increased by a factor of 4. It is desirable to achieve good image quality with fewer captures, since fewer captures reduce imaging system computational power and image sensor readout power consumption [38]. For this novel exposure timing, the SNR can be expressed as

$$SNR(i_{ph}) = \begin{cases} \dfrac{(i_{ph}t_{int})^2}{q(i_{ph}+i_d)t_{int}+\sigma_r^2} & 0 \le i_{ph} \le I_0 - i_d \\ \dfrac{Q_{well}^2}{Q_{well}\cdot q+\sigma_r^2} & I_0 - i_d \le i_{ph} \le 2I_0 - i_d \\ \dfrac{(i_{ph}t_{int}/4^k)^2}{q(i_{ph}+i_d)t_{int}/4^k+\sigma_r^2} & 2^{2k-1}I_0 - i_d \le i_{ph} \le 2^{2k}I_0 - i_d \\ \dfrac{Q_{well}^2}{Q_{well}\cdot q+\sigma_r^2} & 2^{2k}I_0 - i_d \le i_{ph} \le 2^{2k+1}I_0 - i_d \end{cases} \tag{6.7}$$

where $I_0 = \dfrac{Q_{well}}{t_{int}}$, Q_{well} is the well capacity, and t_{int} is the longest integration time. The simulated SNR is plotted in figure 6–14. Compared with a typical multiple sampling technique, this novel multiple sampling enjoys better SNRs, since most photocurrents are able to achieve their full well capacity (see figure 6–15).

To save power, a conventional multiple sampling technique could also increase the exponential capture time by a factor of 4. If so, however, the SNR will dramatically

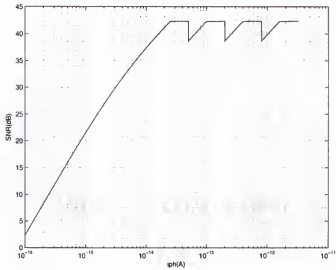


Figure 6–14: SNR of the multiple sampling technique for the TTFS_LM

be degraded. From figure 6–16, we observe numerous 6dB dips in the middle, which is much worse than that of the novel multiple sampling technique.

Of course, this good performance is achieved under the assumption that the integration time can be set arbitrarily long. This assumption, however, does not always hold for DRAMs. The maximum holding time of DRAM sets the longest achievable integration time, i.e., approximately $T_{int} = T_{hold}$. If a SRAM or DRAM with longer holding time is used as local memory, this performance limit will be overcome.

6.4 Contrast Mode TTFS (TTFS_CM)

As noted previously, natural scenes have amplitude spectra that fall inversely with frequency, roughly 1/f [51]. Correspondingly in biological vision systems, the neurons are more sensitive to the local contrast rather than the absolute illuminance value. High contrast is read out ahead of low contrast, which achieves some limited compression. We believe this characteristic is helpful to reduce the collision rate when uniform illuminance is impinged on a sensor. In addition, a contrast-mode image is very useful in the early vision computation.

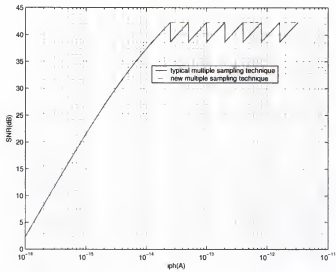


Figure 6–15: Comparison of SNRs when the exposure time in the conventional multiple sampling technique is increased by a factor of 2.

6.4.1 Principles of Current-mode Diffusor

To compute a local spatial average, a resistive network is required [20]. Conventional methods of implementing a resistor in VLSI technology include using a MOSFET or a complex transconductance amplifier. Both methods work in the voltage mode and suffer from a limited range of voltages to demonstrate linear resistance [52]. To circumvent this limitation, a current mode implementation of resistive networks is proposed [53]. Boahen [54] uses the concept of current diffusor illustrated in figure 6–17 to explicitly explain the principle of this novel transistor network.

Assuming that all the PMOS transistors are identical, the transistors connecting to V_R are in the linear region, and the remaining transistors work in the saturation region, then we have

$$\begin{aligned} I_j &= I_0 e^{\frac{(V_{dd}-V_G)+V_{th}}{mU_T}} \left(e^{\frac{V_{dd}-V_j}{U_T}} - e^{\frac{V_{dd}-V_{j+1}}{U_T}} \right) \\ &= I_0 e^{\frac{(V_{dd}-V_G)+V_{th}}{mU_T}} \left(\frac{I_{out_j}}{I_0 e^{\frac{V_{dd}-V_R+V_{th}}{mU_T}}} - \frac{I_{out_{j+1}}}{I_0 e^{\frac{V_{dd}-V_R+V_{th}}{mU_T}}} \right) \\ &= e^{\frac{V_R-V_G}{mU_T}} (I_{out_j} - I_{out_{j+1}}) \end{aligned} \tag{6.8}$$

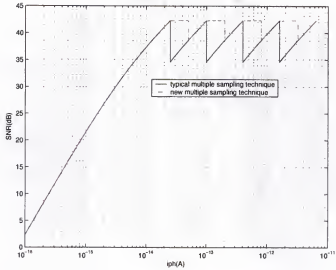


Figure 6–16: Comparison of SNRs when the exposure time in the conventional multiple sampling technique is also exponentially scaled by a factor of 4.

Applying KCL at node V_j, we obtain

$$\begin{aligned} I_{out_j} - I_{in_j} &= I_{j-1} - I_j \\ &= e^{\frac{V_R - V_G}{mU_T}} (I_{out_{j-1}} - 2I_{out_j} + I_{out_{j+1}}) \end{aligned} \tag{6.9}$$

where $(I_{out_{j-1}} - 2I_{out_j} + I_{out_{j+1}})$ is the discrete approximation of the $\frac{d^2}{dx^2}$ operator. Thus, the output current can be modelled as the sum of the input current and the local spatial contrast, which is

$$I_{out_j} = I_{in_j} + I_{contrast_j} \tag{6.10}$$

where $I_{contrast_j} = e^{\frac{V_R - V_G}{mU_T}} (I_{out_{j-1}} - 2I_{out_j} + I_{out_{j+1}})$. Hence, the local contrast information can be achieved after subtracting the input current from the output current.

6.4.2 TTFS_CM Pixel Design

By applying a resistive network, several silicon retinas have been designed to extract the local contrast information. In [55] , an I-V converter is realized by feeding the contrast current through two diode-connected MOSFETs. Since the readout signal is still an analog signal, it is very sensitive to the readout noise. In contrast,

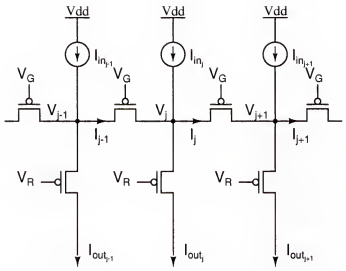


Figure 6–17: Current diffusor circuit

temporal information is believed to be more robust in the noisy environment. Boahen [56] and Barbaro [57] both use an asynchronous readout circuitry to output the temporal information instead of analog values. Like other time-based CMOS image sensors, Boahen's retina also adopts the Pulse Frequency Modulation (PFM) scheme, and too many pulses need to be output for the purpose of good reconstruction. Thus the inevitable high power consumption makes it unfavorable. In [56], the gradient information is obtained by modulating the local contrast current with a predefined sine or cosine wave, which increases the complexity of the front-end circuitry design. To circumvent these disadvantages, we proposed a new contrast-mode imager (TTFS_CM), which still follows the time-to-first-spike readout scheme. The new pixel schematic is shown in figure 6–18 and works as follows:

1. The node X is initially reset to V_{reset}, a middle point of the defined range (V_{ref_low}, V_{ref_high}), by turning on transistor M9.

2. After the $\sim rst$ goes high, a contrast current $I_{contrast_j}$ is generated through the current diffusor and current mirrors formed by *M2-M8*. This generated contrast photocurrent either charges or discharges node X. When the voltage leaves the predefined range, one of the comparators output node flips, and the node req goes high. Then the pixel sends out a request to the row arbiter by pulling down $\sim row\_request(m)$.

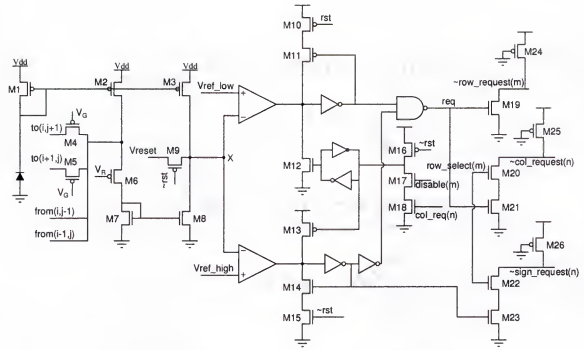


Figure 6–18: Pixel schematic of the TTFS_CM

3. If the row arbiter selects this row by making $\sim row\_select(m)$ high, column request signals $\sim col\_request(n)$ and sign labels $\sim sign\_request(n)$ will be sent to the column latch.

4. After $\sim col_request(n)$s are latched, the corresponding control signals $disable(m)$ and $col\_req(n)$ are generated to disable the pixel by switching on transistor $M12$ and $M13$. Thus the pixel will not fire again until the next reset phase turns off $M12$ and $M13$.

Note that $M10$ and $M15$ are included in the circuitry to ensure the valid logic during the reset period. The comparator can still be a simple 5-transistor opamp.

6.4.3 Simulation Results and Discussion

The proposed circuit was investigated using the SpectreS simulator. After a step signal is fed into a 1-D array, the output signals are measured in terms of firing times and then converted back into currents.

Two different step signals are considered here, and the corresponding simulation results are shown in figure 6–19 and figure 6–20 respectively. From the results, we observe that the current diffusor network successfully extracts the edge information,

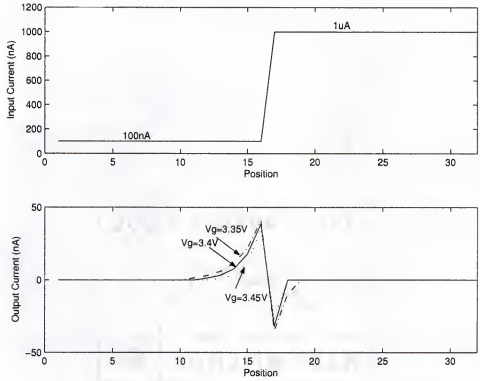


Figure 6–19: Simulation results of the TTFS_CM when input currents are $100nA$ and $1\mu A$.

which is sparsely distributed and may lead to a low collision rate for the asynchronous readout scheme. We also find that the input signal absolute amplitude has an impact on the output signal value, which could be explained as the result of the different initial condition for the differential equation 6.9. Different V_G values are also exploited with a fixed $V_R = 3.4\text{V}$. It shows that when $V_R - V_G$ goes positive, the response output signal is somewhat increased and extended widely. As in [52], we can use *diffusion length* $\lambda = e^{\frac{V_R - V_G}{2mU_T}}$ to explain this phenomenon here. That is with $V_R - V_G$ goes large, λ becomes much significant.

The TTFS_CM actually works in the current mode. In the previous simulation, we didn't consider current mirrors mismatch. If including this effect, the results would be a little "noisy". As we derived in Appendix B, a current mirror working in the subthreshold region has a severe mismatch problem. Since most photocurrents are very weak, we need to size the MOSFETs properly to push them towards working in

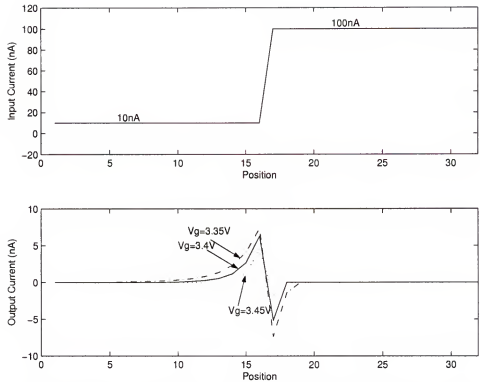


Figure 6–20: Simulation results of the TTFS_CM when input currents are $10nA$ and $100nA$.

the above-threshold region instead of the subthreshold region to lessen the "noisy" mismatch. Considering these factors, the pixel layout is expected to occupy an $80 \times 80\mu m$ area in a $0.6\mu m$ technology.

6.5 Summary and Discussion

In this chapter, several modified TTFS architectures were introduced to solve the potential high collision rate problem associated with a large size TTFS_classic imager. Simulation results have demonstrated the expected performance. Since the TTFS_RS and the TTFS_classic actually share the same idea (a time-based imager with asynchronous readout), and the TTFS_RS is just a minor modification of the TTFS_classic, the TTFS_classic can be easily extended to design the TTFS_RS. The TTFS_CM might be useful in the early vision computation. Edge information is extracted and output using the same asynchronous readout circuitry as in the

TTFS_classic. Once the pixel design is optimized, the TTFS_CM can be easily integrated into the existing TTFS_classic structure. In addition, the complex design of the TTFS_LM is beyond the scope of this dissertation. Therefore, the TTFS_RS, the TTFS_CM and the TTFS_LM are being skipped.

CHAPTER 7
CONCLUSIONS AND FUTURE WORK

In this dissertation, the fundamentals of photodectectors were reviewed. A detailed analysis of conventional CMOS image sensors reveals their DR limitation. To deal with high dynamic range scenes, a novel time-to-first-spike (TTFS) CMOS image sensor is presented to circumvent the DR limitation. The advantages of the TTFS imager include high DR, better SNR, feasibility for video mode applications and low power consumption. A systematic study of an optimal strategy for reference voltage variation in TTFS imagers is also included. The TTFS_classic imager is believed to be a strong competitor among high dynamic range CMOS image sensors, especially for the small to moderate size imager market. The potentially high collision rate, however, limits the TTFS_classic imager from scaling up to a large size. Several modified TTFS architectures are proposed to solve this potentially high collision rate problem.

Due to some practical considerations, a TTFS_classic image was implemented using TSMC $0.18\mu m$ digital CMOS technology in this work. The prototype chip demonstrates the expected performance, which is very exciting and promising. To continue this work, the author believes the following directions need to be considered:

1. The major remaining problem is that of nonuniform motion blur. This is a serious concern for all time-based imagers, particularly in terms of human subjective evaluation of the resulting images. Fundamentally, brighter pixels blur less than darker pixels since brighter pixels are scanned off first. A similar problem occurs with the multiple sampling techniques, however researchers have already begun to address post-processing methods to correct the nonuniform motion blur [37]. We expect that similar post-processing techniques will be developed for TTFS imagers.

2. As we discussed, the TTFS_LM seems to be the best choice to solve the collision problem. Either a SRAM or DRAM must be integrated into each pixel. A small array of the TTFS_LM could be a good example to prove its functionality.

3. An image taken by the TTFS_CM imager is a contrast mode image. A conversion algorithm is demanded to convert a contrast image back to its original image. Currently, our TTFS_CM works in the current mode, whose performance is degraded by current mirrors mismatch. More practical circuit design considerations are needed to get a less "noisy" image.

4. Generally, the sensor response of a time-based imager is nonlinear. So far, all the implemented time-based image sensors are monochromatic sensors, and no color processing is involved. It is believed that a linear sensor response is much beneficial to color processing. Thus, to find a good method dealing with color images for time-based image sensors is another future research direction.

APPENDIX A
DC OFFSET FOR AN OPAMP IN THE WEAK INVERSION REGION

For a differential pair in the strong inversion region, the DC offset voltage is shown in [58] to be

$$V_{OS,in} = \frac{V_{gs} - V_{th}}{2}\left(\frac{-\Delta I_D}{I_D} + \frac{\Delta(W/L)}{(W/L)}\right) - \Delta V_{th} \tag{A.1}$$

It reveals the dependence of $V_{OS,in}$ on device mismatches and bias conditions. For an opamp operating in the weak inversion region, its DC offset demonstrates a similar dependence. However, the explicit DC offset expression is unique due to the different current-voltage characteristic. Assume that both the input transistors and the load resistors in figure A–1 suffer from mismatch, i.e., $V_{th1} = V_{th}$, $V_{th2} = V_{th} + \Delta V_{th}$, $(W/L)_1 = (W/L)$, $(W/L)_2 = (W/L) + \Delta(W/L)$, and $I_{D1} = I_D$, $I_{D2} = I_D + \Delta I_D$. For simplicity, other mismatches are neglected. If all the transistors operate in the weak inversion region and saturate, and $V_{OS,in} = V_{gs1} - V_{gs2}$, we have

$$\begin{aligned} V_{OS,in} &= U_T \ln \frac{I_{D1}}{I_s(W/L)_1} + V_{th1} - \left(U_T \ln \frac{I_{D2}}{I_s(W/L)_2} + V_{th2}\right) \\ &= U_T \left(\ln \frac{I_D}{I_D + \Delta I_D} - \ln \frac{(W/L) + \Delta(W/L)}{(W/L)}\right) - \Delta V_{th} \end{aligned} \tag{A.2}$$

where $I_s = \mu_{eff} C_{ox}(m-1)U_T^2$. Assuming $\Delta I_D/I_D \ll 1$ and $\Delta(W/L)/(W/L) \ll 1$, and noting that for $x \ll 1$ we can approximate $\ln(1+x) \approx x$, we can reduce the above equation to

$$V_{OS,in} = U_T \left(\frac{-\Delta I_D}{I_D} + \frac{\Delta(W/L)}{(W/L)}\right) - \Delta V_{th} \tag{A.3}$$

Compared with equation A.1, the offset for a differential pair in weak inversion is much less than that in strong inversion due to the fact that U_T is typically less than $(V_{gs} - V_{th})/2$ at room temperature.

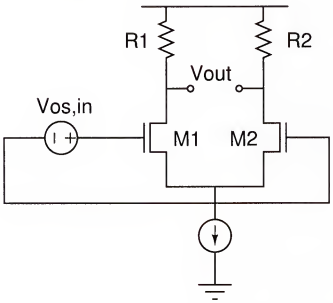


Figure A–1: A differential pair with offset referred to the input

A typical 5-transistor opamp shown in figure A–2 is used in our design. Suppose all the transistors are biased in the weak inversion region and saturated. For the NMOS pair, $V_{OS} = 0$, then we have

$$\frac{\Delta I_N}{I_N} \approx -\frac{\Delta V_{th,N}}{U_T} + \frac{\Delta (W/L)_N}{(W/L)_N} \tag{A.4}$$

With $\frac{\Delta I_P}{I_P} = \frac{\Delta I_N}{I_N}$, we obtain

$$V_{OS,in} = -\Delta V_{th,P} - \Delta V_{th,N} + U_T \left(\frac{\Delta (W/L)_N}{(W/L)_N} - \frac{\Delta (W/L)_P}{(W/L)_P} \right) \tag{A.5}$$

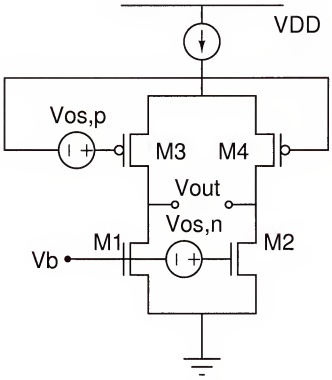


Figure A–2: 5-transistor differential pair with offset referred to the input

Since mismatches are assumed to be independent statistical variables,[1] we can express the standard deviation of the offset as

$$V_{OS,in}^2 = \Delta V_{th,P}^2 + \Delta V_{th,N}^2 + U_T^2 \left(\frac{\Delta(W/L)_N^2}{(W/L)_N^2} + \frac{\Delta(W/L)_P^2}{(W/L)_P^2} \right) \tag{A.6}$$

Here we consider the worst case and simply regard the maximum threshold voltage difference between different corners as the threshold voltage standard deviation, i.e., $\Delta V_{th,P} \approx \Delta V_{th,N} \approx 0.1V$ for TSMC $0.18\mu m$ digital CMOS technology [59], which means the mismatch contribution from ΔV_{th} is more important than that from $\Delta(W/L)$. Then the offset standard deviation will be $V_{OS,in}^2 = 2 \times (0.1V)^2$. If we intend to reduce the offset less than 1.5mV via an autozeroing technique, a finite opamp gain of 100 is required.

[1] We neglect the ΔV_{th} dependence on W here for simplicity

APPENDIX B
CURRENT MIRROR MISMATCH BEHAVIOR

Special care needs to be taken to design current mirrors with small mismatch. The current mirror mismatch can be obtained by calculating the total differential of input current. Since the subthreshold current is $I_D = \mu_{eff} C_{ox}(m-1)U_T^2(W/L)\exp[(V_{gs} - V_{th})/U_T]$, we can estimate the current mirror mismatch as

$$\begin{aligned} \Delta I_D &\approx \frac{\partial I_D}{\partial(W/L)}\Delta\left(\frac{W}{L}\right) + \frac{\partial I_D}{\partial(V_{gs}-V_{th})}\Delta\left(V_{gs}-V_{th}\right) \\ &= I_D\frac{\Delta(W/L)}{W/L} - I_D\frac{\Delta V_{th}}{U_T} \end{aligned} \tag{B.1}$$

where mismatches in $\mu_{eff}C_{ox}(m-1)$ are ignored. It yields

$$\frac{\Delta I_D}{I_D} = \frac{\Delta(W/L)}{W/L} - \frac{\Delta V_{th}}{U_T} \tag{B.2}$$

This result suggests that, to minimize current mismatch, W and L must be *maximized*, and then the threshold mismatch will eventually limit the current mirror performance. In contrast, for a mirror working in the above-threshold region, the current mismatch is given by

$$\frac{\Delta I_D}{I_D} = \frac{\Delta(W/L)}{W/L} - \frac{\Delta V_{th}}{(V_{gs}-V_{th})/2} \tag{B.3}$$

The above equation indicates that we can reduce the current mirror mismatch by pushing the mirror pair into the above-threshold region if $(V_{gs} - V_{th})/2 > U_T$. In our TTFS_CM imager design, we need to intentionally use long channel MOSFETs to ensure the current mirrors operate in the above-threshold region and maximize $(V_{gs} - V_{th})/2$ as well.

REFERENCES

[1] E. R. Fossum, "CMOS image sensors: electronics camera-on-a-chip," *IEEE Transactions on Electron Devices*, vol. 44, no. 10, pp. 1689–1698, 1997.

[2] B. Ackland and A. Dickinson, "Camera on a chip," in *ISSCC Digest of Technical Papers*, 1996, pp. 22–25.

[3] R. Nixon, S. Kemeny, B. Pain, C. Staller, and E. R. Fossum, "256 × 256 CMOS active pixel sensor camera-on-a-chip," *IEEE Journal of Solid-State Circuits*, vol. 31, no. 12, pp. 2046–2050, 1996.

[4] X. Guo, *A time-based asynchronous readout CMOS image sensor*, Ph.D. dissertation, University of Florida, Gainesville, FL, 2002.

[5] M. Shur, *Physics of semiconductor devices*, Prentice Hall, Upper Saddle River, New Jersey, 1990.

[6] A. Gamal, *Introduction to image sensors and digital cameras* (lecture notes), Stanford University, http://www.stanford.edu/class/ee392b, accessed 12/12/2001.

[7] Foveon, *X3 technolgy: direct image sensors*, Foveon, Santa Clara, CA, http://www.foveon.com/X3_tech.html, accessed 06/03/2002.

[8] R. Hornsey, *Design and fabrication of integrated image sensors* (short course notes), University of Waterloo, http://www.cs.yorku.ca/ visor/pdf/CMOS_1.pdf, accessed 05/12/2002.

[9] S. Sze, *Physics of semiconductor devices* (2nd Edition), John Wiley and Sons, New York, 1981.

[10] S. Donati, *Photodetectors: devices, circuits and applications*, Prentice Hall, Upper Saddle River, New Jersey, 2000.

[11] H. Wong, "Technology and device scaling consideratins for CMOS imagers," *IEEE Transactions on Electron Devices*, vol. 43, no. 12, pp. 2131–2142, 1996.

[12] N. V. Loukanova, H. O. Folkerts, J. P. V. Maas, D. W. E. Verbugt, A. J. Mierop, W. Hoekstra, E. Roks, and A. J. P. Theuwissen, "Leakage current modeling of test structures for charaterization of dark current in CMOS image sensors," *IEEE Transactions on Electron Devices*, vol. 50, no. 1, pp. 77 –83, 2003.

[13] I. Inoue, H. Ihara, H. Yamashita, T. Yamaguchi, H. Nozaki, and R. Miyagawa, "Low dark current pinned photo-diode for CMOS image sensor," in *Proceedings of IEEE Workshop on Charged-Coupled Devices and Advanced Image Sensors*, Nagano, Japan, June 1999, pp. 25–28.

[14] H. Cheng and Y. King, "An ultra-low dark current CMOS image sensor cell using n^+ ring reset," *IEEE Electron Device Letters*, vol. 23, no. 9, pp. 538 –540, 2002.

[15] H. Cheng and Y. King, "A CMOS image sensor with dark-current cancellation and dynamic sensitivity operations," *IEEE Transactions on Electron Devices*, vol. 50, no. 1, pp. 91 –95, 2003.

[16] E. R. Fossum, "Active pixel sensors: are CCDs dinosaurs?," in *Proceedings of SPIE Workshop on Charged-Coupled Devices and Advanced Image Sensors*, 1993, vol. 1900, pp. 2–14.

[17] T. Hui, B. Fowler, and A. E. Gamal, "Analysis of temporal noise in CMOS photodiode active pixel sensor," *IEEE Journal of Solid-State Circuits*, vol. 35, no. 1, pp. 92–101, 2001.

[18] A. J. Blanksby and M. J. Loinaz, "Performance analysis of a color CMOS photogate image sensor," *IEEE Transactions on Electron Devices*, vol. 47, no. 1, pp. 55–64, 2000.

[19] O. Yadid-Pecht, "Wide-dynamic-range sensors," *Optical Engineering*, vol. 38, no. 10, pp. 1650–1660, 1999.

[20] C. Mead, *Analog VLSI and Neural Networks*, Addison Wesley, Boston, Massachusetts, 1989.

[21] K. B. S. Decker, R. D. McGrath and C. G. Sodini, "A 256x256 CMOS imaging array with wide dynamic range pixels and column-parrallel digital output," *IEEE Journal of Solid-State Circuits*, vol. 33, no. 12, pp. 2081–2091, 1998.

[22] D. Yang, A. E. Gamal, B. Fowler, and H. Tian, "A 640 x 512 CMOS image sensor with ultrawide dynamic range floating-point pixel-level ADC," *IEEE Journal of Solid-State Circuits*, vol. 34, no. 12, pp. 1821–1834, 1999.

[23] O. Yadid-Pecht and E. R. Fossum, "Wide intrascene dynamic range CMOS APS using dual smapling," *IEEE Transaction on Electron Devices*, vol. 44, no. 10, pp. 1721–1724, 1997.

[24] W.Yang, "A wide-dynamic-range, low-power photosensor array," in *ISSCC Digest of Technical Papers*, 1994, pp. 230–231.

[25] L. McIlrath, "A low-power low-noise ultrawide-dynamic-range CMOS imager with pixel-parral A/D conversion," *IEEE Journal of Solid-State Circuits*, vol. 36, no. 5, pp. 846–853, 2001.

[26] K. Boahen, "A throughput-on-demand address-event transmitter for neuromorphic chips," in *Advanced Research in VLSI*, 1999, pp. 72–86.

[27] E. Culurciello, R. Etienne-Cummings, and K. Boahen, "A biomorphic digital image sensor," *IEEE Journal of Solid-State Circuits*, vol. 38, no. 2, pp. 281–294, 2003.

[28] V. Brajovic and T. Kanade, "A VLSI sorting image sensor: global massively parallel intensity-to-time processing for low-latency adaptive vision," *IEEE Transaction on Robotics and Automation*, vol. 15, no. 1, pp. 67–75, 1999.

[29] Y. Ni, F. Devos, M. Boujrad, and J. H. Guan, "Histogram-equalization-based adaptive image sensor for real-time vision," *IEEE Journal of Solid-State Circuits*, vol. 32, no. 7, pp. 1027–1036, 1997.

[30] D. Stoppa, A. Simoni, L. Gonzo, M. Gottardi, and G. Betta, "Novel CMOS image sensor with a 132-dB dynamic range," *IEEE Journal of Solid-State Circuits*, vol. 37, no. 12, pp. 1846–1852, 2002.

[31] Q. Luo, *Time-based synchronous readout CMOS imager*, Ph.D. dissertation, University of Florida, Gainesville, FL, 2002.

[32] S. Thorpe, D. Fize, and C. Marlot, "Speed of processing in the human visual system," *Nature*, vol. 381, pp. 520–522, 1996.

[33] M. J. Tovee and E. T. Rolls, "Information encoding in short firing rate epochs by single neurons in the primate temporal visual cortex," *Visual Cognition*, vol. 2, no. 1, pp. 35–38, 1995.

[34] J. G. Harris, "The changing roles of analog and digital signal processing in CMOS image sensors," in *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing, 2002*, May 2002, vol. 1, pp. 13–17.

[35] D. Yang and A. E. Gammal, "Comparative analysis of SNR for image sensors with widened dynamic range," in *Proceedings of SPIE*, February 1999, vol. 3649, pp. 197–211.

[36] X. Qi, Q. Luo, X. Guo, and J. G. Harris, "SNR and FPN considerations for time-based wide dynamic range CMOS image sensors," in *Proceedings of World Multconference on Systemics, Cybernetics and Informatics*, July 2001, vol. XX, pp. 66–73.

[37] X. C. Liu and A. E. Gamal, "Synthesis of high dynamic range motion blur free image from multiple captures," *IEEE Transaction on circuits and systems I*, vol. 50, no. 4, pp. 530–539, 2003.

[38] T. Chen and A. E. Gamal, "Optimal scheduling of capture times for a multiple capture imaging system," in *Proceedings of the SPIE Electronic Imaging'2002 Conference*, 2002, vol. 4669, pp. 288–296.

[39] Y. Taur and T. H. Ning, *Fundamentals of modern VLSI devices*, Cambridge University Press, New York, 1998.

[40] C. C. Enz and G. C. Temes, "Circuit techniques for reducing the effects of op-amp imperfections: Autozeroing, correlated double sampling, and chopper stabilization," *Proceedings of the IEEE*, vol. 84, no. 11, pp. 1584–1614, 1996.

[41] B. Razavi, *Principles of data conversion system design*, John Wiley and Sons Inc., New York, 1994.

[42] C. C. Enz, F. Krummenacher, and E. A. Vittoz, "An analytical MOS transistor model valid in all regions of operation and dedicated to low-voltage and low-current applications," *Analog Integrated Circuits Signal Processing*, vol. 8, pp. 83–114, 1995.

[43] B. Beecken and E. Fossum, "Determination of the conversion gain and the accuracy of its measurement for detector elements and arrays," *Applied Optics*, vol. 35, no. 19, pp. 3471–3477, 1996.

[44] K. A. Boahen, "Point-to-point connectivity between neuromorphic chips using address events," *IEEE Transactions on Circuits and Systems II*, vol. 47, no. 5, pp. 416–434, 2000.

[45] P. Larsson, "Resonance and damping in CMOS circuits with on-chip decoupling capacitance," *IEEE Transactions on Circuits and Systems I*, vol. 45, no. 8, pp. 849–858, 1998.

[46] P. Debevec, *Recovering high dynamic range radiance maps from photographs*, [Online.] http://www.debevec.org/Research/HDR/, accessed 09/30/2002.

[47] E. P. Simoncelli and B. A. Olshausen, "Natural image statistics and neural representation," *Annual Review of Neuroscience*, vol. 24, pp. 1193–1216, 2001.

[48] B. A. Wandell, A. E. Gamal, and B. Girod, "Common principles of image acquisition systems and biological vision," *Proceedings of the IEEE*, vol. 90, no. 1, pp. 5–17, January 2002.

[49] S. Kleinfelder, S. Lim, X. Liu, and A. E. Gamal, "A 10,000 frames/s CMOS digital pixel sensor," *IEEE Journal of Solid-State Circuits*, vol. 36, no. 12, pp. 2049–2059, December 2001.

[50] T. Lulé, M. Wagner, M. Verhoeven, H. Keller, and M. Böhm, "100000-pixel, 120-dB imager in TFA technology," *IEEE Journal of Solid-State Circuits*, vol. 35, no. 5, pp. 732–739, May 2000.

[51] D. J. Field, "Relations between the statistics of natural images and the response properties of cortical cells," *Journal of the Optical Society of America*, vol. A, pp. 4.2379–2394, Dec 1987.

[52] S. Liu, J. Kramer, G. Indiveri, T. Delbrück, and R. Douglas, *Analog VLSI: Circuits and Principles*, Cambridge University Press, New York, 1998.

[53] E. A. Vittoz and X. Arreguit, "Linear networks based on transitors," *Electronics Letters*, vol. 29, no. 3, 1984.

[54] K. A. Boahen, *Retinomorphic vision systems: reverse engineering the veterbrate retina*, Ph.D. dissertation, California Institute of Technology, Pasadena, CA, 1989.

[55] J. Shin, D. Park, S. Lee, H. Kim, J. Park, J. Kim, M. Lee, H. Yamada, and H. Yonezu, "A foveated CMOS retina chip for edge detection with local light-adaptation function," in *Proceedings of the 8th International Conference on Neural Information Processing*, Nov. 2001, pp. 14–18.

[56] K. Boahen, "Retinomorphic vision systems I: pixel design," in *proceedings of IEEE International Symposium on Circuits and Systems*, 1996, vol. Supplement, pp. 9–13.

[57] M. Barbaro, P. Burgi, A. Mortara, P. Nussbaum, and F. Heitger, "A 100×100 pixel silicon retina for gradient extraction with steering filter capabilities and temporal output coding," *IEEE Journal of Solid-State Circuits*, vol. 37, no. 2, pp. 160–172, Feb. 2002.

[58] B. Razavi, *Design of analog CMOS integrated circuits*, McGraw-Hill, New York, 2001.

[59] TSMC, *TSMC* $0.18\mu m$ *logic 1P6M silicide* $1.8V/3.3V$ *spice models*, Taiwan Semiconductor Manufacturing Co. LTD, Taiwan, 2002.

BIOGRAPHICAL SKETCH

Xin Qi was born in Shenyang, China. He received the B.S. degree from Northeastern University, China, in 1997 and the M.S. degree from Shanghai Jiaotong University, China, in 2000, both in automatic control. Since 2000, he has been a research assistant in the Computational NeuroEngineering Laboratory (CNEL) at the University of Florida. His research interests include analog and mixed-signal design.

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

John G. Harris, Chair
Associate Professor of Electrical and Computer Engineering

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Robert M. Fox
Associate Professor of Electrical and Computer Engineering

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

William R. Eisenstadt
Associate Professor of Electrical and Computer Engineering

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Stephen S. Eikenberry
Professor of Astronomy

This dissertation was submitted to the Graduate Faculty of the College of Engineering and to the Graduate School and was accepted as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

December 2004

Pramod P. Khargonekar
Dean, College of Engineering

Kenneth J. Gerhardt
Interim Dean, Graduate School